COMPUTER AND CONTROL ENGINEERING

Modality-Agnostic Deep Learning

Funded By	Dipartimento DAUIN FONDAZIONE CRT CASSA DI RISPARMIO DI TORINO [P.iva/CF:06655250014]
Supervisor	CAGLIERO LUCA - luca.cagliero@polito.it
Contact	CAGLIERO LUCA - luca.cagliero@polito.it CHIUSANO SILVIA ANNA - silvia.chiusano@polito.it
Context of the research activity	MultiModal Deep Learning (MMDL) aims at gaining insights into various combinations of media types such as text, audio, and video. Within this research field, a relevant effort has been devoted to extending Deep Natural Language Processing models to handle not only the raw textual documents but also additional data sources (e.g., images). As a drawback, most of the existing MMDL techniques are tailored to specific combinations of input/output data types. Modality-Agnostic Deep Learning (MADL) focuses on relaxing these dependences thus making the devised solutions portable to different scenarios, domains, and tasks. The Ph.D. candidate will study, design, and develop new MADL techniques and will explore their applicability to various real-world contexts.
Objectives	State-of-the-art Deep Learning architectures extract meaning from vary large quantities of data by leveraging their inherent data structure. The data-driven inference process based on Neural Networks relies on simplifying assumptions, called inductive bias, which reflect the underlying data characteristics and, indirectly, their modality of acquisition. MultiModal Deep Learning (MMDL) approaches aim at blending data of multiple data types to overcome the inherent architectural and performance limitations of traditional single-modal solutions. However, their scope is often limited to specific combinations of data types and to particular application contexts. Modality-Agnostic Deep Learning (MADL) focuses on relaxing the dependence of MMDL models on the input/output data types. The purpose is to make the devised solutions more flexible and easily portable to different scenarios, domains, and tasks. Currently, MADL techniques are challenged by - The limited number of existing DL architectures: few attempts to solve the MADL problem have been presented in literature. Hence, the number of unexplored directions is potentially large. - Applications to a rather small number of tasks, domains and applications: compared to the rapidly increasing number of data combinations, tasks, and application domains that have already been explored by the MMDL

	community, still the use of MADL is quite preliminary. - The complexity of domain shift, which requires both cross-modal content adaptation at the source level and the effective handling of unseen domains at the target level. The goal of the Ph.D. research proposal is to deeply explore the extension of existing MMDL techniques towards input or output data type independence, with a particular attention paid to the extension of state-of-the-art Deep Natural Language Processing architectures such Transformers and Seq2Seq models. It also aims at investigating novel approaches to MADL, domain shift, and their application to real-world scenarios, domains, and tasks (e.g., aspect-based sentiment analysis, summarization).

Skills and competencies for the development of the activity	Basic knowledge of machine learning techniques, good programming skills