"Multimodal Machine Learning: A Survey and Taxonomy"
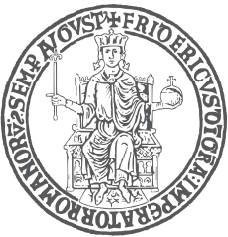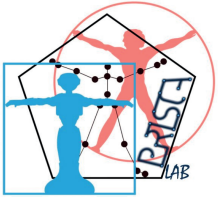
📅 Martedì 16 Giugno 2020   🕐 16:00

📍 Microsoft Teams (codice: qbzoygy)

Durante il prossimo seminario, discuteremo l'articolo di Baltrušaitiset al. di cui sopra, introdotto da **Marco Grazioso**.

Abstract:

Our experience of the world is multimodal - we see objects, hear sounds, feel texture, smell odors, and taste flavors. Modality refers to the way in which something happens or is experienced and a research problem is characterized as multimodal when it includes multiple such modalities. In order for Artificial Intelligence to make progress in understanding the world around us, it needs to be able to interpret such multimodal signals together. Multimodal machine learning aims to build models that can process and relate information from multiple modalities. It is a vibrant multi-disciplinary field of increasing importance and with extraordinary potential. Instead of focusing on specific multimodal applications, this paper surveys the recent advances in multimodal machine learning itself and presents them in a common taxonomy. We go beyond the typical early and late fusion categorization and identify broader challenges that are faced by multimodal machine learning, namely: representation, translation, alignment, fusion, and co-learning. This new taxonomy will enable researchers to better understand the state of the field and identify directions for future research.

PRISCA Lab (Projects of Intelligent Robotics and Advanced Cognitive Systems), Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, Università di Napoli Federico II