

Seminario CMM- Maths&AI

Speaker: Javiera Castillo, Conservatoire National des Arts et Métiers.

Title: "What to align in multimodal contrastive learning?".

Abstract:

Humans perceive the world through multisensory integration, blending the information of different modalities to adapt their behavior. Contrastive learning offers an appealing solution for multimodal self-supervised learning. Indeed, by considering each modality as a different view of the same entity, it learns to align features of different modalities in a shared representation space. However, this approach is intrinsically limited as it only learns shared or redundant information between modalities, while multimodal interactions can arise in other ways. In this work, we introduce CoMM, a Contrastive Multimodal learning strategy that enables the communication between modalities in a single multimodal space. Instead of imposing cross- or intra- modality constraints, we propose to align multimodal representations by maximizing the mutual information between augmented versions of these multimodal features. Our theoretical analysis shows that shared, synergistic and unique terms of information naturally emerge from this formulation, allowing us to estimate multimodal interactions beyond redundancy. We test CoMM both in a controlled and in a series of real-world settings: in the former, we demonstrate that CoMM effectively captures redundant, unique and synergistic information between modalities. In the latter, CoMM learns complex multimodal interactions and achieves state-of-the-art results on the six multimodal benchmarks.

Miércoles 04 de Junio 2025 a las 15:00 hrs.

Link:

<https://uchile.zoom.us/j/98408846527?pwd=5vdhHNTZUw0pOx20eiYsyUE44tmRLM.1>

Sala Maryam Mirzakhani (Multimedia) CMM, Piso 6, Beauchef 851 Edificio Norte.