Guaranteed efficiency and interpretability in self-supervised learning by leveraging tensor methods

Self-supervised learning (SSL) has been one of the main pillars of recent progress in AI and in the so-called foundation models, allowing algorithms to leverage large amounts of unlabeled data (text, image, video, etc.) to learn representations that can later be adapted to different predictive (downstream) tasks [1, 2]. However, despite the practical importance of such approaches, understanding the performance of learned representations is still an open question. This is intimately linked to the learnability of *disentangled* representations [3], including the influence of architecture choices on their stability. In particular, answering the following key questions has direct impact on critical AI applications:

- when can learned representations be interpreted, and what is the influence of architecture and learning method on both their stability and on the performance of downstream tasks, especially on considering specialization over small or heterogeneous datasets?
- how do they perform on heterogeneous (e.g., personalized, or whose statistical distributions is not identical across all datasets) data, which is common in medical applications, in particular, in the framework of federated learning?
- under which conditions can smaller models (with lower carbon impact) perform well on specialized tasks?

Goals: To tackle these challenges, in this project we aim to: 1) study the theoretical aspects of representations constructed by self-supervised (and weakly supervised) learning approaches, which relates to their disentanglement properties, and, based on these theoretical findings, 2) develop new approaches to self-supervised and unsupervised learning guaranteeing stability and interpretability of latent representations under challenging conditions and heterogeneous datasets. Particular interest will be given on developing small models (with low energy and environmental/carbon footprint) to perform in specialized tasks, including, for example, the efficient adaptation of large pretrained models (as in LoRA works [4]), and understanding their interplay and links with large generalist (foundation) models.

These objectives will be achieved by leveraging connections between AI models (e.g., neural networks) and (multilayer) *low-rank tensor decompositions* [5, 6, 7], which supply a rigorous mathematical framework to develop new algorithms and model architectures and to understand their behavior. These techniques will be validated on challenging AI applications, such as the separation of different brain processes in functional magnetic resonance data in a federated learning setting [8], or cross-scene hyperspectral/remote sensing image analysis [9].

Supervision and environment: This research work will be carried out as part of a Ph.D. thesis starting in 2024. The candidate will be jointly supervised by Prof. David Brie, Dr. Ricardo Borsoi and Dr. Konstantin Usevich, members of the Multidimensional Signals (SiMul) team (`https://cran-simul.github.io/`), CRAN Laboratory, University of Lorraine, France.

The candidate will be based in Vandoeuvre-lès-Nancy, France, and work on a dynamic team collaborating with experts on deep learning, tensor decomposition and applications. The position is jointly funded by the CNRS, by the ANR, and by the Grand Est Region within the ENACT AI cluster, with a salary of approximately 2100 euros per month, funding for international travel and access to computing resources that may be necessary during the project.

This PhD offer is provided by the ENACT AI Cluster and its partners. Find all ENACT PhD offers and actions on `https://cluster-ia-enact.ai/`.

Expected profile: Master degree or equivalent, with experience in one or more of the following fields: data analysis, signal processing, machine learning, applied mathematics. A strong mathematical background and good communication skills in English (written and oral) are required.

Candidates should send their application to david.brie@univ-lorraine.fr, ricardo.borsoi@univ-lorraine.fr, konstantin.usevich@univ-lorraine.fr, including an academic CV and a short explanation of their research interests and motivation for this position.

References

[1] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, and J. Tang, "Self-supervised learning: Generative or contrastive," *IEEE transactions on knowledge and data engineering*, vol. 35, no. 1, pp. 857–876, 2021.

[2] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16000–16009, 2022.

[3] X. Wang, H. Chen, Z. Wu, W. Zhu, *et al.*, "Disentangled representation learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

[4] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, *et al.*, "Lora: Low-rank adaptation of large language models.," *ICLR*, vol. 1, no. 2, p. 3, 2022.

[5] R. A. Borsoi, K. Usevich, D. Brie, and T. Adali, "Personalized coupled tensor decomposition for multimodal data fusion: Uniqueness and algorithms," *IEEE Transactions on Signal Processing*, 2024.

[6] Y. Panagakis, J. Kossaifi, G. G. Chrysos, J. Oldfield, M. A. Nicolaou, A. Anandkumar, and S. Zafeiriou, "Tensor methods in computer vision and deep learning," *Proc. IEEE*, vol. 109, no. 5, pp. 863–890, 2021.

[7] A. Cichocki, A.-H. Phan, Q. Zhao, N. Lee, I. Oseledets, *et al.*, "Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives," *Foundations and Trends® in Machine Learning*, vol. 9, no. 6, pp. 431–673, 2017.

[8] B. J. Casey, T. Cannonier, M. I. Conley, A. O. Cohen, D. M. Barch, M. M. Heitzeg, M. E. Soules, T. Teslovich, D. V. Dellarco, H. Garavan, *et al.*, "The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites," *Developmental cognitive neuroscience*, vol. 32, pp. 43–54, 2018.

[9] M. Ye, Y. Qian, J. Zhou, and Y. Y. Tang, "Dictionary learning-based feature-level domain adaptation for cross-scene hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1544–1562, 2017.