Emergence of collective learning in coupled deep neural nets

Lluís Arola-Fernández and Lucas Lacasa

Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC), CSIC-UIB,
Palma de Mallorca, Spain

Keywords— Complex systems, statistical physics, deep learning, federated learning

Unraveling the emergence of collective learning in systems of coupled artificial neural networks points to broader implications for physics, machine learning, neuroscience, and society. The multi-scale nature of the collective learning problem makes it a suitable candidate for a complex systems approach, but this perspective has been largely ignored in the literature in favor of more direct applications and focus on algorithmic performance.

In this poster, I will present the results of our recent preprint [1], where we introduce a minimal model of coupled deep neural networks that condenses several recent decentralized algorithms used in practice –such as federated learning [2], elastic averaging [3] or decentralized Stochastic Gradient Descent [4]– by considering a competition between just two terms: the local learning dynamics in the parameters of each neural network unit, and a diffusive coupling among units that tends to homogenize the parameters of the ensemble. We derive an effective theory for deep linear networks to show that the coarse-grained behavior of our system is equivalent to a deformed Ginzburg-Landau model with quenched disorder [5]. This well stablished framework of statistical physics predicts, in a novel context, depth-dependent disorder-order-disorder phase transitions in the parameters' solutions as we increase the coupling strength between the neural units. These phase transitions reveal a depth-delayed onset of a collective learning phase –when units learn from each other– and a low-rank microscopic learning path. We validate our coarse-grained theory in coupled ensembles of realistic neural networks trained on the MNIST and CIFAR-10 datasets under privacy constraints. Interestingly, our experiments confirm that individual neural networks –trained on private data– can fully generalize to unseen data classes when the collective learning phase emerges.

This work unveils the physics of collective learning and contributes to the mechanistic interpretability of deep learning in decentralized settings. By applying concepts and tools from statistical physics to current distributed machine learning systems, our proof of concept illustrates how a complex system's approach, that leverage coarse-graining ideas and learning scales separation, can unveil the key mechanisms of interacting deep learning models. In an era where artificial intelligence is advancing rapidly, understanding these interactions is crucial, especially considering the profound and as-yet-unknown impacts that these 'black-box' deep learning algorithms may have on society.

References

[1] L. Arola-Fernández and L. Lacasa. An effective theory of collective deep learning. arXiv:2310.12802, 2023.

[2] H. B. McMahan and others. Communication-Efficient Learning of Deep Networks from Decentralized Data. In *Proceedings of PMLR*, pages 1273–1282, 2017.

[3] S. Zhang, A. E. Choromanska, Y. LeCun. Deep learning with elastic averaging SGD. In *Adv. in NIPS*, 28, 2015.

[4] J. Wang and G. Joshi. Cooperative SGD: A Unified Framework for the Design and Analysis of Local-Update SGD Algorithms. *J. of M. L. Research*, 22(213):1–50, 2021.

[5] P. Hohenberg and A. Krekhov. An introduction to the Ginzburg-Landau theory of phase transitions and nonequilibrium patterns. *Phys. Rep.*, 572, 1-42, 2014.