

Speaker: Alex SHEVCHENKO

Title: Theory of Compression via Shallow Autoencoders

Abstract: Autoencoders are a popular model in many branches of machine learning and lossy data compression. However, their fundamental limits, the performance of gradient methods and the features learnt during optimization remain poorly understood, even in the two-layer setting. In fact, earlier work has considered either linear autoencoders or specific training regimes (leading to vanishing or diverging compression rates). We address this gap by focusing on non-linear two-layer autoencoders trained in the challenging proportional regime in which the input dimension scales linearly with the size of the representation. Our results characterize the minimizers of the population risk, and show that such minimizers are achieved by gradient methods; their structure is also unveiled, thus leading to a concise description of the features obtained via training. For the special case of a sign activation function, our analysis establishes the fundamental limits for the lossy compression of Gaussian sources via (shallow) autoencoders. Finally, while the results are proved for Gaussian data, numerical simulations on standard datasets display the universality of the theoretical predictions in the shallow case.