

STATISTICAL PHYSICS OF DEEP LEARNING

Pietro Rotondo

University of Parma, Italy

ABSTRACT

Understanding how and why Deep Learning works is considered a major challenge in Mathematics, Computer Science and Theoretical Physics. In this talk, I will present recent advances in Deep Learning Theory, focusing on the so-called proportional regime (which is formally defined as the thermodynamic limit where both the size of the trainset and the size of the hidden layers are taken to infinity keeping their ratio finite).

I will show that this regime is suitable to rationalise the behaviour of realistic finite width deep neural networks. In particular, I will show that an effective statistical physics description based on “kernel renormalization” captures feature learning in finite architectures and identifies different forms of representation learning that strongly depend on the topology and weight-sharing properties of the architecture.