Is Evolution an Algorithm? Effects of local entropy in unsupervised learning and protein evolution

Doctoral Dissertation

Doctoral Program in Physics (34.th cycle)

Is evolution an algorithm?

Effects of local entropy in unsupervised learning and protein evolution

Matteo Negri

Synthesis of Doctoral Thesis

In the present thesis I study artificial neural networks with the lens of statistical mechanics, field that has proven very useful in understanding the structure of the solution space of combinatorial optimization problems such as the *perceptron*, the basic unit of any artificial neural network.

The rise of artificial neural networks in the last decade has been staggering and now deep (i.e. many-layered) neural networks applications are ubiquitous in technology and science. Despite this fact, a comprehensive theory that explains this success is still missing: according to the classical statistical learning framework, deep networks should *not* have good generalization properties because they are utilized in the over-parametrized regime (i.e. much more parameters than examples), where statistical learning theory predicts overfitting for any class of models.

An interesting conjecture which has emerged in various contexts argues that the flatness of the minima can lead to good generalization in the over-parametrized regime. For this reason recently a theory has been developed that connects the generalization capabilities of artificial neural networks with the geometrical properties of the error loss functions that is minimized for learning. This theory makes use of the concept of *local entropy*, a function that counts the number of other solutions around a given solution. So far theoretical results on generalization have been limited to the basic *supervised learning* setting of the *teacher-student* model. Other empirical results are available which are not limited to a teacher-student setting, but they are still examples of supervised learning.

Supervised learning means training an artificial neural network to do some classification task. This means finding a rule to fix the parameters given a set of example patterns, in a way that the network assigns the correct label to each example. The adjective supervised refers to the fact that each example pattern must be provided with a label.

The *teacher-student model* is the prototypical classification problem. In this setting the examples are independent and identically distributed randomly-generated patterns and the labels are provided as the output of a second network (called teacher) that is randomly initialized and that has the same architecture of the first network (called student).

The main goal of this thesis is to explore the effect of local-entropy-inspired algorithms in situations that are progressively more different from the teacher-student scenario. First I study a perceptron model on a gaussian-mixture data distribution, then I switch to an unsupervised setting (i.e. the examples are not labeled) and finally I study a system that is completely unrelated to neural networks where local entropy proves to be useful to understand the evolution of proteins.

I found that looking for wide flat minima in neural networks is relevant even when the problem is convex while the naive optimization of the loss is not sufficient to generalize well.

I also showed that wide solutions of sparse autoencoders implement representation of data with different properties than typical solutions, and I observe that these representations are closer to the actual structure of data when such structure is known. I explored some real-world applications of this improved feature extraction on amino-acid sequences of homologous proteins families.

Finally, I abandoned neural networks and applied the concept of local entropy to the problem of protein folding. In particular, I disentangled the study of the sequence from that of the three-dimensional structure and showed how structures that maximize local entropy seem to be better suited from the point of view of evolutionary fitness.