Internship 2018

Un-normalized Distribution

Supervisor: Aurélien Decelle & Michèle Sebag
Labs: TAU – CNRS – INRIA – Paris-Sud
Mail: decelle, sebag@lri.fr
URL: `http://www.lri.fr/~sebag/Stages/Stage_NoZ_19.pdf`

Abstract :
Identifying the joint distribution underlying a finite sample set has many applications, ranging from generative learning to feature selection. Among the most prominent approaches are Restricted Boltzman Machines [?], defined as a bipartite graph involving visible features v_i and hidden features h_j, and characterizing a distribution $\hat{P}$ with:

$$\hat{P}(v_1, \dots v_D) = \frac{1}{Z} \sum_{(h_1,\dots,h_d)} exp\left(- \sum_{i,j} W_{i,j} h_i v_j - \sum_i A_i v_i - \sum_j B_j H_j \right)$$

where Z is the normalization term ensuring that P actually is a probability (summing to 1), involving a double summation on the v and the h. A main source of difficulty for training an RBM is to take into account the Z term in medium to large-dimensional spaces. Contrastive divergence [?] and its many variants aim to address this difficulty.

Goal

The goal of the internship is to to train an RBM (i.e. identify parameters W, A and B) and side-step the computation of the Z factor, considering two settings: i) the target distribution P is known; ii) an iid sample generated from target P, is available.

Proposed approach

The proposed approach takes inspiration from Energy-based learning [?] and adversarial learning [?]. The conjecture is that learning a function $f : X \mapsto \mathbb{R}^+$ monotonous w.r.t. P (i.e. $P(x) > P(x') + \eta \Rightarrow f(x) > f(x') + \epsilon$) will provide the required information to sample the target distribution.

The challenge will be to define *informative negative examples*. Random samples will not be informative enough as the target distribution expectedly lie on a manifold. The proposed approach will consider "chimeras", obtained by combining existing samples such that i) the chimeras are close to the real samples; ii) it is likely that they are not on the manifold.

References

[1]]A Tutorial on Energy-Based Learning - Yann LeCun 2006.
[2] Ian J. Goodfellow: NIPS 2016 Tutorial: Generative Adversarial Networks.