Proposition de stage 2022/2023

Restricted Coulomb Machines

THEME: Machine learning and statistical physics
RESPONSABLE: Cyril Furtlehner
LABORATOIRE: TAU team INRIA Saclay, LISN
ADRESSE: Université Paris-Saclay, 91405 Orsay Cedex
MAIL: Cyril.Furtlehner@inria.fr

1 Learning Generative models

Recently deep neural networks have experienced a spectacular development. They are nowadays commonly used for voice recognition in smartphones or for automatic handwriting recognition. While these spectacular applications results from solving supervised learning tasks, some progress have been as well observed in parallel with the same kind of architecture in the non-supervised context. In that case the goal is mainly to build a probability distribution function (pdf) of the data, which in turn can be used as a generative model, or twisted to solve classification tasks. There is a large variety of generative models [1] with various degree of sophistication. Associated learning algorithms are generally based on log likelihood ascent. The restricted Boltzmann machine (RBM) corresponds to a family of generated models which realizes a good trade-off between efficiency and interpretability. With a simple architecture, it presents good performances for learning a pdf, while allowing for theoretical analysis of its behavior and interpretation of its features. Still in many cases the training remains unsatisfactory. With the lenses of statistical physics[3] a rather comprehensive picture of the RBM as been developped in the last few years. In addition a convex relaxation of the RBM has been proposed in [2]. In this model -call it Restricted Coulomb Machine (RCM)-, the usual sigmoid activation function is decomposed via a the 1-d Coulomb potential leading to a continuous set of features. For data laying on a flat low dimensional space this leads to an exact training of the distribution of features.

2 Objectives of the internship

We would like to extend the feasibility of the RCM to real data cases not necessarily restricted to low dimensional cases. Given a data set, this can be possibly done following these steps:

- Generate a set of features, with various possible methods (e.g. pre-training of an ordinary RBM or e.g. via self supervised learning of the data).
- Decompose and replicate the features into "Coulomb" features by designing a statistical selection procedure.
- Find via convex optimization the optimal weights associated to the selected features.

Some refinement or constraints could be added to this procedure like imposing symmetries (e.g. translation) or specific architecture like corresponding to source models. Various experiments are expected to be conducted, first on synthetic datasets where the performance of the method can be assessed and then on some real data like population genetic sequences or solar space weather time series related to ongoing projects in the lab.

Further reading

[1] R.R. Salakhutdinov, "Learning Deep Generative Models", Annual Review of Statistics and its Applications (2015) — http://www.cs.toronto.edu/∼rsalakhu/papers/annrev.pdf

[2] Decelle, Furtlehner, "Exact Training of Restricted Boltzmann Machines on Intrinsically Low Dimensional Data", Phys.Rev.Lett. (2021)

[3] Decelle, Furtlehner, "Restricted Boltzmann Machines, recent advances and Mean Field Theory", Ch.Phys. B (2021)