9th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2021)


Contribution ID: **193** Type: **Sectional reports**

# STATISTICAL MECHANICS APPROACH FOR DEEP-BELIEF NEURAL NETWORKS EXPLORATION

Humans and other animals can understand concepts from only a few examples. while standard machine learning algorithms require a large number of examples to extract hidden features. Unsupervised learning is procedure of revealing hidden features from unlabeled data.
In deep neural network training, unsupervised data pre-training increases the final accuracy of the algorithm by decreasing an initial parameter space from which fine-tuning begins. However, there are few theoretical papers devoted to detail unsupervised learning description. Crucial reason is that unsupervised learning process in deep neural network is usually complicated. That's why understanding the mechanism of it in elementary models plays an important role.
Boltzmann machines are the basic unit for developing deep-belief networks. Due to their ability to reveal hidden internal representations and solve complex combinatorial problems, they are used in machine learning and statistical patterns infer. Boltzmann machines are neural networks with symmetrically connected layers, divided into two categories – visible and hidden. In this work we consider a Restricted Boltzmann machine (RBM), with links between neurons of different layers, but without internal.
To solve computational problems, the machine firstly pass training, where its parameters – neuron activation thresholds $\theta$ and weights on the edges $\xi$, are stochastically changed according to the selected algorithms. After that, the visible layer is initialized with a given state, and the system evolves to a stationary distribution. Finally, the output layer represents the solution of the problem.
Dealing with deep networks often issues the loss of interpretation of the obtained features, i.e., the loss of physical essence.
Despite success in practical applications, the rigorous mathematical description of Boltzmann machines remains a challenge. In studies, the coefficients of weights on edges are considered fixed, and their distribution is extracted in training. The study of RBM's can be done with statistical mechanics, which development was contributed by the famous Soviet scientist N.N. Bogoliubov. The symmetry property of the weights matrix and the equality of the main diagonal determine the similarity of the Boltzmann machine with the physical model of spin glasses. RBM with binary bonds is equivalent to a bipartite spin glass with layer variables of different nature. The visible layer consists of binary Ising spins and the hidden layer consists of real Gaussian spins.
The purpose of this work is to physically describe the OMB and study its process modes by analytical and numerical methods.

This work was supported by the RFBR No. 18-29-10014. This research has been supported by the Interdisciplinary Scientific and Educational School of Moscow University «Photonic and Quantum Technologies. Digital Medicine».

## Summary

**Primary authors:** PEREPELKIN, Evgeny (JINR); POLYAKOVA, Rimma (LIT JINR); Dr INOZEMTSEVA, N.; SADOVNIKOV, B. (Faculty of Physics, Lomonosov Moscow State University); RUDAMENKO, R. (Faculty of Physics, Lomonosov Moscow State University)

**Presenter:** RUDAMENKO, R. (Faculty of Physics, Lomonosov Moscow State University)

**Session Classification:** Distributed computing, HPC and ML for solving applied tasks