

POSTER PRESENTATION

Open Access

Multilayer perceptrons, Hopfield's associative memories, and restricted Boltzmann machines

Shin-ichi Asakawa

From The Twenty Third Annual Computational Neuroscience Meeting: CNS*2014
Québec City, Canada. 26-31 July 2014

This study was intended to describe multilayer perceptrons (MLP), Hopfield's associative memories (HAM), and restricted Boltzmann machines (RBM) from a unified point of view. Despite of mutual relation between three models, for example, RBMs have been utilizing to construct deeper architectures than shallower MLPs. The energy function in HAM is analogous to the Ising model in statistical mechanics, and it connects microscopic physics to thermodynamics. The canonical partition function Z in the Boltzmann distribution is also utilized RBMs. Asynchronous updating and contrastive divergence (CD) based upon Gibbs sampling is also related. Therefore, it seems to be worth considering these three models within a common framework. This attempt might lead to "one algorithm hypothesis.", which insists that our brains might rule a single but universal rule.

An algorithm, which someone could find out in a region, may be applicable to other regions. Multilayer perceptrons (henceforth, MLP) are feed forward models for pattern recognition and classification. Hopfield proposed another kind of neural network models for associative memory and optimization (HAM). Hinton adopted the restricted Boltzmann machines (RBM) in "Deep Learning" in order to construct deeper layered neural networks. The energy employed in RBMs are elicited the generalized EM

algorithm, which was closely related to the energy employed by HAM. In spite of other various differences, see Table 1, it is worth considering to compare among them. At least, an attempt is worth attempting to explain all of them in a unified terminology.

HAM and RBM have symmetrically weighted connections, $w_{ij} = w_{ji}$, although generalized Boltzmann machines can not satisfy this constraints. Similarly, there are no feedback connections in MLP in general. When we denote a connection weight from j -th unit to i -th unit as w_{ij} , $w_{ij} \in \mathbb{R}$, $w_{ji} = 0$ in MLP. When we consider a merged weight matrix W , all the models can be considered as identical.

The construction methods adopted by Deep Learning are based upon RBMs. One of key concepts to success for constructing multilayer deep architecture is the non-linearity, because units in hidden layer in RBMs are binary. The non-linearity seems to play an important role to construct deep architecture. When we suppose to abandon CD and binary feature, multilayer architecture might replace one weight matrix $W = W_1 W_2 \dots W_p$. Also, we can consider a thought experiment with only one hidden unit in RBM. If $h = 0$, then there are no meanings at all. If $h = 1$, then it must be an identity mapping, or at least, it might be extract the eigenvector vector corresponded to the maximum eigenvalue value in data matrix X . This

Table 1 Summary of MLP, HAM, and RBM

	Architecture	Learning	Flow Direction	Within Layer Connection	Update
MLP	Layered	Gradient Descent	Feed Forward	No	Synchronous
HAM	Mutual	No (EM Algorithm)	No Directions	Yes	Asynchronous
RBM	Layered	Contrastive Divergence	Mutual	No	Synchronous

Correspondence: asakawa@ieee.org
Center for Information Sciences, Tokyo Woman's Christian University, 2-6-1
Zempukoji, Suginami-ku, Tokyo 167-8585, Japan



© 2014 Asakawa; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

might be equivalent to the algorithm proposed by Oja (1985). Since Deep Learning architecture network models trained via RBMs have no within layer connections, we might not be able to reject a possibility that a hidden unit hi might be trained to detect exactly the same features as another hidden unit hj . In order to avoid these situations, we must prepare a large number of binary hidden units more than the entropy being involved in input data set. RBM has no assumptions about within layer connections, it might success to detect important features among data matrix via CD. However, this constraint might weaken slightly, when we would introduce the EM algorithm to be estimate the states of latent variables, and an online algorithm of HAM. This might bring us to an idea “semi restricted Boltzmann machines.”

Published: 21 July 2014

doi:10.1186/1471-2202-15-S1-P223

Cite this article as: Asakawa: Multilayer perceptrons, Hopfield's associative memories, and restricted Boltzmann machines. *BMC Neuroscience* 2014 15(Suppl 1):P223.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

