

Accepted Manuscript

Learning in the machine: Random backpropagation and the deep learning channel

Pierre Baldi, Peter Sadowski, Zhiqin Lu

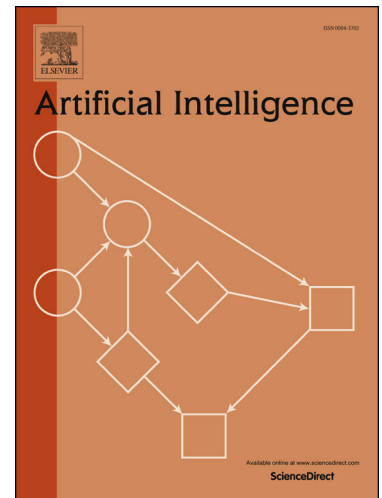
PII: S0004-3702(18)30098-5
DOI: <https://doi.org/10.1016/j.artint.2018.03.003>
Reference: ARTINT 3061

To appear in: *Artificial Intelligence*

Received date: 7 December 2016
Revised date: 21 December 2017
Accepted date: 15 March 2018

Please cite this article in press as: P. Baldi et al., Learning in the machine: Random backpropagation and the deep learning channel, *Artif. Intell.* (2018), <https://doi.org/10.1016/j.artint.2018.03.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Learning in the Machine: Random Backpropagation and the Deep Learning Channel

Pierre Baldi^{1,*}, Peter Sadowski¹, and Zhiqin Lu²

Abstract

Abstract: Random backpropagation (RBP) is a variant of the backpropagation algorithm for training neural networks, where the transpose of the forward matrices are replaced by fixed random matrices in the calculation of the weight updates. It is remarkable both because of its effectiveness, in spite of using random matrices to communicate error information, and because it completely removes the taxing requirement of maintaining symmetric weights in a physical neural system. To better understand random backpropagation, we first connect it to the notions of local learning and learning channels. Through this connection, we derive several alternatives to RBP, including skipped RBP (SRPB), adaptive RBP (ARBPA), sparse RBP, and their combinations (e.g. ASRBP) and analyze their computational complexity. We then study their behavior through simulations using the MNIST and CIFAR-10 benchmark datasets. These simulations show that most of these variants work robustly, almost as well as backpropagation, and that multiplication by the derivatives of the activation functions is important. As a follow-up, we study also the low-end of the number of bits required to communicate error information over the learning channel. We then provide partial intuitive explanations for some of the remarkable properties of RBP and its variations. Finally, we prove several mathematical results, including the convergence to fixed points of linear chains of arbitrary length, the

*Corresponding author.¹ Department of Computer Science, University of California, Irvine. ² Department of Mathematics, University of California, Irvine.

Download English Version:

<https://daneshyari.com/en/article/6853010>

Download Persian Version:

<https://daneshyari.com/article/6853010>

[Daneshyari.com](https://daneshyari.com)