Modern Hopfield Networks

Sepp Hochreiter
Johannes Kepler University
Linz, A-4040 Linz, Austria
hochreit@bioinf.jku.at

Abstract

We propose a new paradigm for deep learning by equipping each layer of a deep learning architecture with modern Hopfield networks. The new paradigm is a new powerful concept comprising functionalities like pooling, memory, and attention for each layer. Associative memories date back to the 1960/70s and became popular through Hopfield Networks in 1982. Recently, we saw a renaissance of Hopfield Networks, the modern Hopfield Networks, with a tremendously increased storage capacity and an extremely fast convergence. We generalize modern Hopfield Networks with exponential storage capacity to continuous patterns. Their update rule ensures global convergence to local energy minima and they converge in one update step with exponentially low error. Surprisingly, the transformer attention mechanism is equal to the update rule of our new modern Hopfield Network with continuous states. The new modern Hopfield network can be integrated into deep learning architectures as layers to allow the storage of and access to raw input data, intermediate results, or learned prototypes. These Hopfield layers enable new ways of deep learning, beyond fully-connected, convolutional, or recurrent networks, and provide pooling, memory, association, and attention mechanisms. We demonstrate the broad applicability of the Hopfield layers across various domains. Hopfield layers improved state-of-the-art on three out of four considered multiple instance learning problems as well as on immune repertoire classification with several hundreds of thousands of instances. On the UCI benchmark collections of small classification tasks, where deep learning methods typically struggle, Hopfield layers yielded a new state-of-the-art when compared to different machine learning methods. Finally, Hopfield layers achieved state-of-the-art on two drug design datasets.

References

[1]   **github:** https://github.com/ml-jku/hopfield-layers
[2]   **arXiv:** "Hopfield Networks is All You Need", https://arxiv.org/abs/2008.02217
[3]   **blog:** https://ml-jku.github.io/hopfield-layers/
[4]   **Videos:** "Modern Hopfield Networks" (S. Hochreiter) https://youtu.be/bsdPZJKOlQs
[5]   "Hopfield Networks in 2021" (S. Hochreiter & D. Krotov) https://youtu.be/k3YmWrK6wxo
[6]   "Performers & Memory" (S. Hochreiter & K. Choromanski) https://youtu.be/BuDhitKZ_YY