

Abstract

Dr. Eric Vanden-Eijnden basic research project proposes a radically different approach to the mathematics of neural networks, through the analogy with the dynamics and statistics of an ensemble of interacting particles

The performance of neural networks on high-dimensional data sets suggests that it may be possible to represent high-dimensional functions with controllably small errors, potentially outperforming standard interpolation methods such as Galerkin truncation or finite element that have been the workhorses of scientific computing but suffer from the curse of dimensionality. This project pro-poses and exploits a theoretical framework to justify these observations and put learning with neural networks on firm mathematical foundations. This is achieved by mapping the parameters of a neural network to a system of particles relaxing with an interaction potential determined by the loss function used to train the network. This analogy enables us to use the powerful mathe-matical tools developed for interacting particle systems to analyze the behavior of the empirical distribution of these parameters / particles. The approach shows that the loss landscape becomes asymptotically convex at the level of the particle / parameter distribution. This permits a rederiva-tion of the universal approximation theorem for neural networks. It additionally shows that the optimal representation can be achieved through stochastic gradient descent (SGD), the algorithm ubiquitously used for parameter optimization in machine learning. The approach also indicates that, for a network of size n , the fluctuations around the optimal representation arise at a scale $O(n^{-1})$, for suitable choices of the batch size, and the error prefactor can be obtained by solving an explicit equation involving the network kernel. This offers the possibility for the first time to adjust the network architecture to minimize the representation error. The research proposed here will not only have theoretical and practical implications for the ways neural networks are currently used in ML, but will also greatly extend the range of applicability of ML by marrying it with scientific computing to perform high-dimensional calculations out of reach nowadays.