POLITECNICO DI TORINO

DISMA Dipartimento di Scienze Matematiche G. L. Lagrange

ECCELLENZA 2018 · 2022

Online seminar

Wednesday December 02, 2020 at 17:30
Hosted on: Zoom

Eric Vanden-Eijnden

Professor at the Courant Institute of Mathematical Sciences, New York University

Trainability and accuracy of artificial neural networks

Prof. De Gregorio introduces the seminar.

Abstract

The methods and models of machine learning (ML) are rapidly becoming *de facto* tools for the analysis and interpretation of large data sets. Complex classification tasks such as speech and image recognition, automatic translation, decision making, etc. that were out of reach a decade ago are now routinely performed by computers with a high degree of reliability using (deep) neural networks. These performances suggest that DNNs may approximate high-dimensional functions with controllably small errors, potentially outperforming standard interpolation methods based e.g. on Galerkin truncation or finite elements that have been the workhorses of scientific computing.

In support of this prospect, in this talk Prof. Vanden-Eijnden will present results about the trainability and accuracy of neural networks, obtained by mapping the parameters of the network to a system of interacting particles relaxing on a potential determined by the loss function. This mapping can be used to prove a dynamical variant of the universal approximation theorem showing that the optimal neural network representation can be attained by (stochastic) gradient descent, with a approximation error scaling as the inverse of the network size. He will also show how these findings can be used to accelerate the training of networks and optimize their architecture, using e.g nonlocal transport involving birth/death processes in parameter space.

Biography

Eric Vanden-Eijnden earned his doctorate in 1997 from the Université Libre de Bruxelles under the supervision of Radu Bălescu. Since 2003, he is a professor of mathematics at the Courant Institute of Mathematical Sciences, New York University. He also held visiting position at Cambridge University, the Institute for Advanced Study, and the University of California, Berkeley.

In 2009 he was awarded the Germund Dahlquist Prize of the Society for Industrial and Applied Mathematics "for his work in developing mathematical tools and numerical methods for the analysis of dynamical systems that are both stochastic and multiscale" and in 2011 he won SIAM’s J.D. Crawford Prize for outstanding research in nonlinear science.