Journal of Machine Learning

Journal of Machine Learning (JML) publishes high quality research papers in all areas of machine learning, including innovative algorithms of machine learning, theories of machine learning, important applications of machine learning in AI, natural sciences,social sciences, and engineering etc. The journal emphasizes a balanced coverage of both theory and practice. The journal is published in a timely fashion in electronic form. All articles in JML are open-access and there is no charge for the authors.

Beyond the Quadratic Approximation:

The Multiscale Structure of Neural Network Loss Landscapes

Chao Ma, Daniel Kunin, LeiWu, Lexing Ying

DOI: 10.4208/jml.220404, J. Mach. Learn., 1 (2022), pp. 247-267.

Communicated by: Zhi-Qin John Xu

Category: Theory

Summary for general readers:

The loss landscape of neural networks cannot be studied only using local quadratic approximation. We examined the loss landscape in a region larger than the reach of good quadratic approximation. Numerically, we observe that neural network loss functions possess a multiscale structure, manifested in two ways: (1) in a neighborhood of minima, the loss mixes a continuum of scales and grows subquadratically, and (2) in a larger region, the loss shows several separate scales clearly. Using the subquadratic growth, we are able to explain the Edge of Stability phenomenon observed for the gradient descent (GD) method. Using the separate scales, we explain the working mechanism of learning rate decay by simple examples. Finally, we study the origin of the multiscale structure and propose that the non-convexity of the models and the non-uniformity of training data is one of the causes. By constructing a two-layer neural network problem we show that training data with different magnitudes give rise to different scales of the loss function, producing subquadratic growth and multiple separate scales.

Sponsored by the Center for Machine Learning Research, Peking University & AI for Science Institute, Beijing.

Global Science Press https://www.global-sci.org/jml