

CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - May 22, 2020

Double-descent phenomenon in modern machine learning

Lu Lu

While breakthroughs in machine learning and artificial intelligence are changing society, our fundamental understanding has lagged behind. It is traditionally believed that fitting models to the training data exactly is to be avoided as it leads to poor performance on unseen data. However, powerful modern classifiers frequently have near-perfect fit in training, a disconnect that spurred recent intensive research and controversy on whether theory provides practical insights. In this work, we show how classical theory and modern practice can be reconciled within a single unified performance curve and propose a mechanism underlying its emergence. We believe this previously unknown pattern connecting the structure and performance of learning architectures will help shape design and understanding of learning algorithms.

Other relevant papers will also be presented.

1. Belkin et al., PNAS, 2019.
2. Nakkiran et al., ICLR, 2019.
3. Bos & Opper, NIPS, 1997.
4. Advani & Saxe, arXiv:1710.03667, 2017.
5. Spigler et al., J. Phys. A, 2019.
6. Neal et al., arXiv:1810.08591, 2018.