



Canadian Association
of Physicists

Association canadienne
des physiciens et physiciennes

Contribution ID: 17

Type: Oral (Non-Student) / Orale (non-étudiant(e))

Sufficient is better than optimal for training neural networks

Tuesday 10 June 2025 17:15 (15 minutes)

As the 2024 Nobel Prizes highlight, neural networks have emerged both as an important physical paradigm, and for representing complex phenomena beyond physics. The fidelity of neural network representations of complex phenomena relies on a fundamental tension between universality and generalizability. In this talk, we will argue that, for neural network training, optimization is the enemy of generalizability. We will show that reframing neural network training in physical terms opens new paths to generalizable networks. We will describe an alternative set of training algorithms that exploit the mathematical physics of filters. Using arguments from information geometry, we will show that filter-based training algorithms yield a set of “sufficient training” methods that outperform optimal training methods, e.g., Adam. We will show that sufficient training can be used to “retrofit” networks that were overfit by optimal training. We will give examples where sufficient training improves generalizability when deployed from the outset. We will describe an open-source implementation of sufficient training we term “simmering”. Using these results we will make the case that maintaining physical perspectives on neural networks is pivotal for their continued application to complex phenomena in physics and beyond.

Keyword-1

Neural networks

Keyword-2

Information geometry

Keyword-3

Authors: BABAYAN, Irina (Queen’s University); ALIAHMADI, Hazhir (Queens University); Prof. VAN ANDERS, Greg (Queen’s University)

Presenter: Prof. VAN ANDERS, Greg (Queen’s University)

Session Classification: (DTP) T3-10 Particle Physics and Field Theory | Physique des particules et théorie des champs(DPT)

Track Classification: Technical Sessions / Sessions techniques: Theoretical Physics / Physique théorique (DTP-DPT)