

**The Weizmann Institute of Science
Faculty of Mathematics and Computer Science**

Machine Learning and Statistics Seminar

Room 1, Ziskind Building
on Wednesday, Jun 19, 2024
at 11:15

Yair Carmon
TAU

will speak on

Making SGD as Parameter-Free as Possible

Abstract:

While stochastic optimization methods drive continual improvements in machine learning, choosing the optimization parameters—particularly the learning rate (LR)—remains challenging. In this talk, I will describe our work on eliminating LR tuning from stochastic gradient descent (SGD) under convexity assumptions. Our starting point is a novel post-hoc empirical certificate for the SGD step size choice, which yields strong parameter-free guarantees via a bisection procedure. This certificate also inspires a tuning-free dynamic SGD step size formula, which we call Distance over Gradients (DoG). For smooth stochastic objectives, we combine DoG with UniXGrad (Kavis et al., 2019) to obtain the first accelerated parameter-free method. Finally, we develop a “price of adaptivity” framework that allows us to evaluate the inherent cost of not knowing problem parameters in advance. In several settings, our lower bounds nearly match existing upper bounds, establishing there is no parameter-free lunch.

Joint work with Maor Ivgi, Itai Kreisler, and Oliver Hinder.