SPRING 2021
ECE RISING STAR
SEMINAR SERIES

NICOLAS
LOIZOU

LARGE-SCALE OPTIMIZATION FOR MACHINE LEARNING AND DATA SCIENCE

Mila - Quebec Artificial
Intelligence Institute

ABSTRACT
Stochastic gradient descent (SGD) is the workhorse for training modern large-scale supervised machine learning models. In this talk, we will discuss recent developments in the convergence analysis of SGD and propose efficient and practical variants for faster convergence. We will start by presenting a general yet simple theoretical analysis describing the convergence of SGD under the arbitrary sampling paradigm. The proposed analysis describes the convergence of an infinite array of variants of SGD, each of which is associated with a specific probability law governing the data selection rule used to form minibatches. The result holds under the weakest possible assumptions providing for the first time the best combination of step-size and optimal minibatch size. We will also present a novel adaptive (no-tuning needed) learning rate for SGD. We will introduce a stochastic variant of the classical Polyak step-size (Polyak, 1987) commonly used in the subgradient method and explain why the proposed stochastic Polyak step-size (SPS) is an attractive choice for setting the learning rate for SGD. We will provide theoretical convergence guarantees for the new method in different settings, including strongly convex, convex, and non-convex functions, and demonstrate the strong performance of SGD with SPS compared to state-of-the-art optimization methods when training over-parameterized models. Finally, we will close with a brief presentation of how standard optimization methods can also solve smooth games (min-max optimization problems) through the Hamiltonian viewpoint.

BIO
Nicolas Loizou is a Postdoctoral Research Fellow at Mila - Quebec Artificial Intelligence Institute and the Université de Montréal. He completed his Ph.D. studies in Optimization and Operational Research at the University of Edinburgh, School of Mathematics, in 2019. Prior to that, he received his undergraduate degree in Mathematics from the National and Kapodistrian University of Athens in 2014, and in 2015 obtained his M.Sc. degree in Computing from Imperial College London. During the fall of 2018, he was a research intern at Facebook AI Research, Montreal, Canada. His research interests include large-scale optimization, machine learning, randomized numerical linear algebra, distributed and decentralized algorithms, game theory, and deep learning. His current research focuses on the theory and applications of convex and non-convex optimization in large-scale machine learning and data science problems. He has received several awards and fellowships, including OR Society's 2019 Doctoral Award (runner-up) for the "Most Distinguished Body of Research leading to the Award of a Doctorate in the field of Operational Research" and the IVADO Postdoctoral Fellowship.

DETAILS
Tuesday, February 16, 2021 at 9:30 am CST.

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON