



THE WEIZMANN INSTITUTE OF SCIENCE
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 1 ,Ziskind Building
on Wednesday, Jun 19, 2019
at 11:15

Daniel Soudry
Technion

Theoretical and Empirical Investigation of Several Common Practices in Deep
Learning

Abstract:

We examine several empirical and theoretical results on the training of deep networks. For example, Why are common "over-fitting" indicators (e.g., very low training error, high validation loss) misleading? Why, sometimes, early-stopping time never arrives? Why can adaptive rate methods (e.g., adam) degrade generalization? Why commonly used loss functions exhibit better generalization than others? Why use weight decay before batch-norm? When can we use low numerical precision, and how low can we get? and discuss the practical implications of these results.

Bio == Since October 2017, Daniel soudry is an assistant professor (Taub Fellow) in the Department of Electrical Engineering at the Technion, working in the areas of machine learning and theoretical neuroscience. Before that, he did his post-doc (as a Gruss Lipper fellow) working with Prof. Liam Paninski in the Department of Statistics, the Center for Theoretical Neuroscience the Grossman Center for Statistics of the Mind at Columbia University. He did his Ph.D. in the Department of Electrical Engineering at the Technion, Israel Institute of technology, under the guidance of Prof. Ron Meir. He received his B.Sc. degree in Electrical Engineering and Physics from the Technion.