**Title: Distributed Training of Deep Neural Networks**

Modern deep neural networks are comprised of Billions of parameters, which require massive amounts of data and time to train. Exponential growth of these networks along the years has made it impractical to train them from scratch on a single GPU/machine. Distributing the computations over several GPUs/machines can drastically reduce this training time, however, stochastic gradient descent (SGD), which is typically used to train these networks, is an inherently sequential algorithm. As a result, training deep neural networks on multiple workers is difficult, especially when using non-dedicated cloud resources trying to maintain high efficiency, scalability and final accuracy. I this talk we will survey some of the new ideas in this scope and discuss their potential.

**Bio:**

Prof. Assaf Schuster of the Computer Science Department is the head of the new AI center at the Technion. He is a Fellow of the ACM and the IEEE, with more than 200 published papers in highly selected venues. His interests and publications are in the wide scope of distributed and scalable data mining, big and streaming data technologies including management, analytics & prediction, cyber security and system/IoT vulnerabilities, privacy preserving, cloud resource management and more. He consulted leading hi-tech companies and participated in the bumpy journey of several startups, two of which he co-founded. More information: https://assaf.net.technion.ac.il/