

## *ML@GT Seminar: Alireza Aghasi*

Wednesday, March 14th

12pm in Marcus Nano Technology, room 1116

---



Alireza Aghasi is currently an assistant professor in the Institute for Insight at the Robinson College of Business. Prior to this he was a research scientist with the IBM T.J. Watson research center, Yorktown Heights. From 2015 to 2016 he was a postdoctoral associate with the computational imaging group at the Massachusetts Institute of Technology, and between 2012 and 2015 he served as a postdoctoral research scientist with the compressed sensing group at Georgia Tech. His research fundamentally focuses on optimization theory and statistics, with applications to various areas of data science, artificial intelligence, modern signal processing and physics-based inverse problems.

Title: Pruning Deep Neural Networks with Net-Trim: Deep Learning and Compressed Sensing Meet

Abstract: We introduce and analyze a new technique for model reduction in deep neural networks. Our algorithm prunes (sparsifies) a trained network layer-wise, removing connections at each layer by addressing a convex problem. We present both parallel and cascade versions of the algorithm along with the mathematical analysis of the consistency between the initial network and the retrained model. We also discuss an ADMM implementation of Net-Trim, easily applicable to large scale problems. In terms of the sample complexity, we present a general result that holds for any layer within a network using rectified linear units as the activation. If a layer taking inputs of size  $N$  can be described using a maximum number of  $s$  non-zero weights per node, under some mild assumptions on the input covariance matrix, we show that these weights can be learned from  $O(s \log N/s)$  samples.

Please RSVP [HERE](#) for this lecture as food will be served