Avner May

535 West 112th Street, Apt 404 • New York, NY 10025 • (301)518-5058
avnermay@cs.columbia.edu

EDUCATION

Columbia University — **New York, NY**
PhD Candidate in Computer Science — ***Sept. 2011 - Present***
MS in Computer Science — ***May 2012***
GPA: 4.07/4.00
Advisor: Michael Collins
Honors: Recipient of the Department Chair's Distinguished Fellowship
Teaching: Course Assistant for "Computer Networks" and "Challenges in Cloud and Mobile Computing"
Relevant Courses: Machine Learning, Adv. Machine Learning, Statistical Inference

Harvard University — **Cambridge, MA**
Bachelor of Arts in Mathematics, Secondary Field Computer Science — ***June 2009***
Cumulative GPA: 3.60/4.00 – Mathematics GPA: 3.58/4.00 – Computer Science GPA: 3.86/4.00
Honors: Certificate of Distinction in Teaching (Spring 2008).
Relevant Courses: Intro. to CS I/II, Theory of Computation, Data Structures & Algorithms, Efficient Algorithms, Probability Theory

Charles E. Smith Jewish Day School — **Rockville, MD**
GPA: 4.54/4.00 (highest in graduating class) — ***February 2005***
Honors (selection):
The Keter Torah Award for excellence in all academic disciplines
The Samuel W. Greenhouse Award for excellence in Mathematics
Rensselaer Medal for excellence in Math and Science

RESEARCH EXPERIENCE AT COLUMBIA — 2011-PRESENT

Acoustic Modeling (Sept. 2013-Present): My main research areas are in Machine Learning and Speech Recognition. Specifically, I have been working on acoustic modeling, which is a large-scale multi-class classification problem important in speech recognition systems. Most recently, I have focused on scaling kernel methods to this domain using approximation techniques. Although kernel methods have been extensively studied, and have a strong theoretical foundation, there has been relatively little success scaling these methods to large-scale problems like speech recognition and computer vision. Currently, deep neural networks (DNNs) are the state of the art in these domains. I am interested in better understanding the differences between these two families of models, as well as improving these methods.

Social Network Analysis (2011-2013): I researched whether social networks like Facebook or Twitter are efficient systems for delivering content of interest to their users. This work had two parts. A theoretical component, in which I used game theoretic tools to prove, given my social network model, that these systems are quite efficient in equilibrium; and an empirical part, in which I analyzed data from social networks to validate the assumptions and predictions of the model.

PUBLICATIONS

Compact Kernel Models for Acoustic Modeling via Random Feature Selection
A. May, M. Collins, D. Hsu, B. Kingsbury
ICASSP 2016

A Comparison Between Deep Neural Nets and Kernel Acoustic Models for Speech Recognition
Z. Lu, D. Guo, A.B. Garakani, K. Liu, **A. May**, A. Bellet, L. Fan, M. Collins, B. Kingsbury, M. Picheny, F. Sha
ICASSP 2016

How to Scale Up Kernel Methods to Be As Good As Deep Neural Nets
A. May, Z. Lu, K. Lu, A. Garakani, D. Guo, A. Bellet, L. Fan, F. Sha, M. Collins, B. Kingsbury
arXiv:1411.4000, 2014

Filter & follow: How social media foster content curation
A. May, A. Chaintreau, N. Korula, S. Lattanzi
SIGMETRICS 2014

WORKSHOP PRESENTATIONS

NIPS 2015: Feature Extraction Workshop, Montreal, Canada, Dec. 11, 2015.
SANE 2015 (Speech and Audio in the Northeast), New York, NY, Oct. 22, 2015.
NYML 2015 (New York Machine Learning Symposium), New York, NY, Mar. 13, 2015.
EC 2013: Workshop on Social Computing and User Generated Content, Philadelphia, PA, June 16, 2013.
New York Computer Science and Economics Day (NYCE), New York, NY, Dec. 3rd 2012.
Interdisciplinary Workshop on Information and Decision in Social Networks (WIDS), Cambridge, MA, Nov. 8-9, 2012.

WORK EXPERIENCE

Google Research – Large Scale Machine Learning Research Group — **New York, NY**
Research Intern — ***Summer 2015***
Worked on model compression, a research area which attempts to train more compact models in the case where larger more powerful models already exist. Worked on developing better methods for model compression. Performed experiments using Torch.

Microsoft Research – Speech and Dialogue Research Group — **Redmond, WA**
Research Intern — ***Summer 2014***
Worked on training acoustic models from the raw speech signal. Specifically, was interested in seeing whether it was possible to train the matrices which perform the fourier transform and mel-binning, as part of the classical MFCC feature extraction pipeline. Performed extensive experiments with, and made large improvements to, the Computational Network Toolkit (CNTK), an open-source C++ machine learning toolkit developed by MSR.

Microsoft Corporation – Windows Communication Foundation (WCF) — **Redmond, WA**
Software Development Engineer — ***Aug. 2009 - July 2011***
Developer on the Messaging Framework Team. Designed and implemented features to facilitate the development of distributed applications.
Honors: Received "Gold Star Bonus Award" for contributions to team.

Microsoft Corporation – Windows Workflow Foundation (WF) — **Redmond, WA**
Software Development Engineer Intern — ***Summer 2008***
Designed and implemented program for validating Windows Workflow programs. Integrated it with Microsoft Visual Studio.

Harvard University – Mathematics Department — **Cambridge, MA**
Course Assistant for Math 23b: Linear Algebra and Real Analysis II — ***Spring 2008***
Led weekly review of material covered in class. Held weekly office hours.
Honors: Awarded "Certificate of Distinction in Teaching" based on evaluations by students.

University of Maryland – Granular Physics Lab — **College Park, MD**
Research Assistant — ***Summer 2007***
Conducted research in granular physics with Professor Wolfgang Losert. Studied the propagation of avalanches in excitable media using the tools of image processing. Programmed extensively in IDL (interactive data language).

The Inter-American Development Bank (IDB) — **Washington, DC**
Knowledge Intern — ***Summer 2006***
Worked as part of a team in the Development Effectiveness and Strategic Planning Department to revamp IDB's Project Alert Identification System (PAIS). Created a strategic proposal with recommendations for improving this system.

Math Tutor — **Washington, DC**
Self-Employed — ***2004 - 2009***
Helped students develop and strengthen their mathematical skills through regular mentoring sessions.

SKILLS

Computer: Matlab, Torch, Java, C#, C, C++, CUDA
Language: ***Spanish:*** Native speaker. **Hebrew**: Proficient.