AMIT SHAVIT, PHD

http://shavitamit.github.io • Boston, MA • (781) 696-8282 • shavitamit@gmail.com

OBJECTIVE

Experienced data scientist with strong communication skills. Passionate about leadership roles, client-facing roles, and technical roles involving machine learning and quantitative finance.

EDUCATION

Ph.D. in Chemical and Biomolecular Engineering *University of Pennsylvania* (Philadelphia, PA) *June 2015*

B.S. in Chemical Engineering *University of Massachusetts Amherst* (Amherst, MA) *May 2010*

- Graduated *summa cum laude* in Chemical Engineering with a minor in Chemistry • **GPA:** 3.81

RELEVANT EXPERIENCE

Thomson Reuters – *Data Scientist* at *Thomson Reuters Labs*, Boston, MA *April 2016 – Present*

- Investigated cutting-edge text analysis and natural language processing techniques for machine learning and text classification purposes
- Researched deep learning methodologies and architectures in Python; utilized vowpal wabbit for advanced text/feature generation
- Implemented an ensemble of deep learning and NLP methods that classified news and its significance to stock portfolios; found ways to maximize precision/recall/accuracy using ten-fold cross validation and grid search

Thomson Reuters – *Technology Associate (Rotational Program)* *August 2015 – April 2016*

TR Legal Rotation (New York, NY): *Redefining Legal Service Delivery Models Using Big Data Technologies*

- Researched cutting-edge machine learning algorithms for classification and association of legal documents for integration into a new product in TR Legal
- Implemented three classifiers (Naïve Bayes, k-Nearest Neighbors, cosine similarity) using Java to facilitate accurate legal content classification, and saved content to a MongoDB database
- Designed a web app using Python Flask/MongoDB/HTML/Javascript to present the classified content and allow users (with logins) to write individualized comments and edits that get saved in the database

Collaboration with R&D (New York, NY): *Social Network Research and Analysis, Quantitative Finance*

- Researched and developed quantitative finance algorithms to analyze the social network "StockTwits"
- Utilized big data technologies such as Hive and Python Pandas to facilitate analysis of 20 GB of tweet data
- Manuscript detailing the results of this study published in *DSAA 2016* (acceptance rate 21%)

University of Pennsylvania – *PhD Candidate* under Rob Riggleman, Philadelphia, PA *September 2010 – June 2015*

- Developed algorithms and code to analyze confined glassy polymers using high performance computing and big data technologies (*e.g.,* C & C++, Python, Unix, parallelization, supercomputers/clusters)
- Published five first-author papers in reputable journals; presented research in 17 local and national venues

Academia Sinica – *National Science Foundation EAPSI Fellow*, Taipei, Taiwan *June 2013 – August 2013*

- Developed CUDA-based Brownian Dynamics simulation code; worked in a team; presented results to NSF

SELECTED AWARDS

Audience Favorite Talk (*U. Penn 2014*) • Best Poster (*U. Penn 2013*) • NSF EAPSI Research Fellow (*NSF 2013*)

SELECTED LEADERSHIP ACTIVITIES

Graduate Student Symposium – Co-president, University of Pennsylvania *2014*

- Co-organized 2014 symposium; number of attendees increased more than 50% over previous year

Excellence in Teaching Award – University of Pennsylvania Center of Teaching and Learning *Sept. 2012 – Dec. 2012*

- Developed teaching philosophy; discussed methods for engaging students; received feedback on teaching style

PROJECTS

I actively participate in collaborative Kaggle competitions and have finished in the top 5.5% and top 7% on competitions involving ML, NLP, ensemble methods, as well as image classification using deep neural networks

SKILLS

C & C++, Python, Linux, BASH, Front End Stack (html/css/js), MATLAB, Parallel Computing, LaTeX, git, Java, Hive, SQL, MongoDB, API Development, and MS Office Suite

Citizen of the United States of America