

Amanda J. Minnich

amandajeann119@gmail.com
amandaminnich.info
linked.in/amandajeannminnich

Research Interests

Applied machine learning researcher focused on the adversarial and platform health space. At Twitter I use machine learning to detect spam and abuse campaigns. Previously I led the Molecular Data-Driven Modeling Team at Lawrence Livermore National Laboratory.

Education

- **UNIVERSITY OF NEW MEXICO** **Albuquerque, NM**
Ph.D., Computer Science, GPA 4.04/4.0
2014 – 2017
Dissertation title: “Spam, Fraud, and Bots: Improving the Integrity of Online Social Media Data”
- **UNIVERSITY OF NEW MEXICO** **Albuquerque, NM**
M.S., Computer Science, GPA: 4.06/4.0
2011 – 2014
- **UNIVERSITY OF CALIFORNIA, BERKELEY** **Berkeley, CA**
B.A., Integrative Biology, 3.66/4.0
2005 – 2009

Technical Skills

- **Programming Languages, Libraries, and Tools**
Python (including Pandas, sklearn, TensorFlow, Matplotlib, etc.) (10 YOE), SQL (BigQuery, PostgreSQL, MySQL, Presto) (8 YOE), Git (9 YOE), Docker and Kubernetes (2 YOE)
- **Machine Learning/Data Science Methods**
Experience in supervised and unsupervised algorithms, both classical ML and deep learning, various types of feature selection/pruning, hyperparameter optimization, etc.

Professional Experience

- **TWITTER INC.** **San Francisco, CA**
Data Scientist, Scaled Enforcement Heuristics Team *Jan 2020 - Present*
I create automated pipelines to detect inauthentic coordinated behavior using unsupervised machine learning methods. My work spans the full spectrum of research, prototyping, A/B testing, and productionization, as well as firefighting high-priority spam and abuse issues on the platform.
- **LAWRENCE LIVERMORE NATIONAL LABORATORY** **Livermore, CA**
Machine Learning Research Scientist, Molecular Data-Driven Modeling Team Lead *July 2017 - Jan 2020*
I served as the data-driven modeling tech lead for the ATOM Consortium, where we integrated machine learning into the drug discovery process. I was the chief architect for the ATOM Modeling PipeLine (AMPL), an open source deep learning pipeline, which supports the whole machine learning life cycle: data processing; feature extraction/normalization; model training and evaluation; ad hoc prediction generation; and model/data storage, provenance, and validation.
- **GROUPON INC.** **Palo Alto, CA**
Data Science Intern *Summer 2015*
Designed a predictive bid regression model with an expanded feature set for improved SEM ad performance and implemented smart keyword generation for products using NLP analysis of product descriptions.
- **MANDIANT, A FIREEYE COMPANY** **Albuquerque, NM**
Data Science Research Intern *Summer 2014*
Wrote malware family random forest classifier that was put into production and is currently part of company's toolkit and modified JavaScript's D3 library's Force Layout to implement a Barnes-Hut approximation of t-SNE.
- **CENTER FOR CYBERDEFENDERS, SANDIA NATIONAL LABORATORY** **Albuquerque, NM**
Data Science Research Intern *Summer 2013*
Applied k-means clustering to Frobenius norm inter-year distances for dimension reduction of system call trace-based Markov chain matrices and created random forest classifier model to identify malware.

Publications

Conferences

- [C1] **Amanda J. Minnich**, N. Chavoshi, D. Koutra, and A. Mueen. Botwalk: Efficient adaptive exploration of twitter bot networks. In *ASONAM*, pages 467–474. ACM, 2017. **17.2% Acceptance Rate.**
- [C2] N. Abu-El-Rub, **Amanda J. Minnich**, and A. Mueen. Impact of referral incentives on mobile app reviews. In *ICWE*, pages 351–359. Springer, 2017. **28% Acceptance Rate.**
- [C3] N. Abu-El-Rub, **Amanda J. Minnich**, and A. Mueen. Anomalous reviews owing to referral incentive. In *ASONAM*, pages 313–316. ACM, 2017. **25% Acceptance Rate.**
- [C4] **Amanda J. Minnich**, N. Abu-El-Rub, M. Gokhale, R. Minnich, and A. Mueen. Clearview: Data cleaning for online review mining. In *ASONAM*, pages 555–558. IEEE Press, 2016. **13% Acceptance Rate.**
- [C5] A. Mueen, N. Chavoshi, N. Abu-El-Rub, H. Hamooni, and **Amanda J. Minnich**. Awarp: fast warping distance for sparse time series. In *ICDM*, pages 350–359. IEEE, 2016. **8.6% Acceptance Rate.**
- [C6] **Amanda J. Minnich**, N. Chavoshi, A. Mueen, S. Luan, and Mi. Faloutsos. Trueview: Harnessing the power of multiple review sites. In *WWW*, pages 787–797, 2015. **14.1% Acceptance Rate.**
- [C7] M. Lakin, **Amanda J. Minnich**, T. Lane, and D. Stefanovic. Towards a biomolecular learning machine. In *International Conference on Unconventional Computing and Natural Computation*, pages 152–163. Springer, Berlin, Heidelberg, 2012.

Journals

- [J1] K. McLoughlin, C. Jeong, T. Sweitzer, **Amanda J. Minnich**, M. Tse, B. Bennion, J. Allen, S. Calad-Thomson, T. Rush, and J. Brase. Machine learning models to predict inhibition of the bile salt export pump. *Journal of Chemical Information and Modeling*, 61(2):587–602, 2021. PMID: 33502191.
- [J2] **Amanda J. Minnich**, K. McLoughlin, M. Tse, J. Deng, A. Weber, N. Murad, B. Madej, B. Ramsundar, T. Rush, S. Calad-Thomson, J. Brase, and J. Allen. Ampl: A data-driven modeling pipeline for drug discovery. *Journal of Chemical Information and Modeling*, 60(4):1955–1968, 2020. PMID: 32243153.
- [J3] N. Murad, K. Pasikanti, B. Madej, **Amanda J. Minnich**, J. McComas, S. Crouch, J. Polli, and A. Weber. Predicting volume of distribution in humans: Performance of in silico methods for a large set of structurally diverse clinical compounds. *Drug Metabolism and Disposition*, 2020.
- [J4] A. Mueen, N. Chavoshi, N. Abu-El-Rub, H. Hamooni, **Amanda J. Minnich**, and J. MacCarthy. Speeding up dynamic time warping distance for sparse time series data. *Knowledge and Information Systems*, 54(1):237–263, 2018.
- [J5] **Amanda J. Minnich**. Spam, fraud, and bots: Improving the integrity of online social media data (PhD Dissertation). 2017.
- [J6] M. Lakin, **Amanda J. Minnich**, T. Lane, and D. Stefanovic. Design of a biochemical circuit motif for learning linear functions. *Journal of the Royal Society Interface*, 11(101):20140902, 2014.

Service to Profession and Awards

- Artificial Intelligence Track Co-Chair, Grace Hopper Celebration (2019 and 2020)
- Co-Organizer, Fifth Computational Approaches for Cancer Workshop at SC (2019)
- Program Committee Member, KDD19, CSoNet19, ASONAM17, ASONAM18, and ASONAM19
- President and Co-founder, UNM Women in Computing (2015-2017)
- Grace Hopper Celebration Scholar, for Outstanding Women in Computer Science (2014)
- NIH Programs in Biology and Biomedical Sciences Fellow (2013-2015)
- National Science Foundation Graduate Research Fellow (2012-2017)

Invited Talks & Unpublished Presentations

- [P1] **Amanda J. Minnich.** Safety, reproducibility, performance: Accelerating cancer drug discovery with ML and HPC technologies. CompBioMed, 2019.
- [P2] **Amanda J. Minnich,** N. Chavoshi, and A. Mueen. Taming social bots: Detection, exploration and measurement. SIAM International Conference on Data Mining, 2019.
- [P3] **Amanda J. Minnich.** Using GPUs to generate reproducible workflows to accelerate drug discovery. GPU Technology Conference, 2019.
- [P4] **Amanda J. Minnich.** Safety, reproducibility, performance: Accelerating cancer drug discovery with ML and HPC technologies. HPC User Forum, 2019.
- [P5] **Amanda J. Minnich.** Safety, reproducibility, performance: Accelerating cancer drug discovery with cloud, ML, and HPC technologies. Fourth Computational Approaches for Cancer Workshop (CAFCW18) at SC, 2018.
- [P6] **Amanda J. Minnich.** Utilizing container technology to streamline data science. National Laboratories Information Technology Summit, 2018.
- [P7] **Amanda J. Minnich** on behalf of Nikan Chavoshi. Temporal patterns in bot activities. 7th Temporal Web Analytics Workshop at WWW, 2017.