DAVID S. ROSENBERG

515 E 72nd St Apt 16K • New York, NY 10021
www.linkedin.com/in/dr-david-rosenberg
(510) 541-3897 • david.davidr@gmail.com

EDUCATION

2002–2008 **University of California, Berkeley**, Berkeley, CA
Ph.D., Statistics
Research: machine learning theory, natural language processing

2000–2002 **Harvard University**, Cambridge, MA
S.M., Applied Mathematics (Computer Science focus)

1995–2000 **Yale University**, New Haven, CT
B.S., Mathematics, *cum laude*, with Distinction in the Major

1998–1999 **Technical University of Budapest**, Budapest, Hungary
The Budapest Semesters in Mathematics study abroad program

EXPERIENCE

Bloomberg, New York, NY

Feb 2022–Present *Head of Machine Learning Strategy Team, CTO Office*

- Lead team of senior data scientists with mandate to collaborate with engineering on technical and strategic projects, conduct machine learning research, spread knowledge internally through education and consulting, maintain collaborations with experts in academia and industry, and communicate Bloomberg's achievements and expertise externally
- Guide Bloomberg's strategic investment in machine learning

Sep 2015–Jan 2022* *Machine Learning Architect, CTO Office*

- Led internal research efforts on extracting data from images, leading to 3 peer-reviewed publications
- Consulted internally on wide range of machine learning and statistical challenges (e.g. large language models, anomaly detection, network security, causal inference)
- Enabled engineers to use state-of-the-art deep learning models by advocating for and overseeing Bloomberg's first large-scale investment in graphics processing units (GPUs)
- Conceived, proposed, and managed development of Bloomberg's data science platform, which centralized building of machine learning solutions and greatly increased efficiency
- Created and led ML EDU, Bloomberg's internal machine learning education program (9 courses, 500+ students), which has significantly increased understanding and use of machine learning at Bloomberg; publicly released Foundations of Machine Learning course, which has over 100,000 views

Hawkfish, New York, NY

Mar 2020–Dec 2020 *Team Lead, Research and Development*

- Led team of data scientists on improving Hawkfish's prediction model for the 2020 US presidential election; developed neural network model that outperformed internal traditional models as well as other available industry predictions (e.g. FiveThirtyEight and The Economist) in the general election

* From Nov 2019–Dec 2020, worked at Hawkfish as a Bloomberg contractor; experience listed separately

- Diagnosed sources of error in Florida election forecast using ecological regression
- Designed and built framework to automate the building and evaluation of machine learning models; used framework to investigate large number of alternative data sources to supplement standard demographic profiles (e.g. TV viewership, historical donation data, news readership)

Nov 2019–Feb 2020 — *Senior Data Scientist*

- Developed approach to making reliable support predictions when important demographic information is missing (an issue for as many as 20% of voters)
- Developed sampling method to substantially reduce compute time required to forecast large number of congressional district-level election outcomes
- Developed vote reassignment scheme to play out hypothetical increases in vote share for particular candidates

Jan 2015–Present — **New York University**, New York, NY
Adjunct Associate Professor, Center for Data Science

- Created and taught DS-GA 3001: Tools and Techniques for Machines Learning, an advanced Masters-level course covering an unusual blend of statistics and machine learning topics, chosen for their relevance to applied business settings (2021)
- Designed curriculum and taught DS-GA 1003: Machine Learning, the core machine learning course for the M.S. in Data Science program; received the Center for Data Science's *Professor of the Year Award* in 2015 and 2016 (2015–2019)

Jan 2014–Aug 2015 — **YP (Formerly Yellow Pages)**, New York, NY
Chief Scientist, YP Mobile Labs

- Led team of data scientists in design, implementation, and deployment of automated real-time ad-bidding system, which was responsible for executing all mobile advertising campaigns
- Designed novel statistical methods to optimize bidding strategies used for ad buying

Sense Networks (acquired by YP), New York, NY
Feb 2012–Dec 2013 — *Chief Scientist*
Aug 2011–Feb 2012 — *Lead Scientist*
Sep 2008–Aug 2011 — *Research Scientist*

- Led research and development in location data analysis, ad targeting, and real-time bidding (RTB) strategies

Jun 2007–Dec 2010 — **Discovereads (acquired by Goodreads)**, San Francisco, CA
Scientific Advisor

- Consulted on several problems related to book recommendations and finding "similar books"

Dec 2005–Sep 2007 — **Aptima**, Woburn, MA
Statistical Consultant

- Consulted on natural language processing problems related to internet chat rooms, including conversation thread separation and message-type classification

Apr 2005–Aug 2005 — **Zillow.com**, Seattle, WA
Statistical Consultant

- Consulted on techniques for real estate valuation, including non-parametric regression

The MITRE Corporation, Bedford, MA

Aug 2000–Aug 2002 *Signals Analyst*

- Analyzed techniques for inverting cryptographic one-way functions; used generative and discriminative techniques to classify helicopter radar signatures; applied restless-arm bandit framework to radar scheduling problem

PAPERS

2021 "Dual Reinforcement-Based Specification Generation for Image De-Rendering"
R. Pasunuru, **D. Rosenberg**, G. Mann, M. Bansal
Workshop on Scientific Document Understanding at AAAI, 2021

2019 "Challenges in End-to-End Neural Scientific Table Recognition"
Y. Deng, **D. Rosenberg**, G. Mann
International Conference on Document Analysis and Recognition (ICDAR), 2019

2019 "Improving Grey-Box Fuzzing by Modeling Program Behavior"
S. Karamcheti, G. Mann, and **D. Rosenberg**
Workshop on Machine Learning for Software Engineering (ML4SE), 2019

2019 "Visual attention model for cross-sectional stock return prediction and end-to-end multimodal market representation learning"
R. Zhao, Y. Deng, M. Dredze, A. Verma, **D. Rosenberg**, A. Stent
Proceedings of the Florida AI Research Symposium (FLAIR), 2019

2018 "Adaptive Grey-Box Fuzz-Testing with Thompson Sampling"
S. Karamcheti, G. Mann, and **D. Rosenberg**
11th ACM Workshop on Artificial Intelligence and Security (AISec), 2018

2017 "Scatteract: Automated Extraction of Data from Scatter Plots"
M. Cliche, **D. Rosenberg**, D. Madeka, C. Yee
Machine Learning and Knowledge Discovery in Databases (ECML PKDD), 2017

2015 "Collaborative Place Models"
B. Kapicioglu, **D. Rosenberg**, R. Schapire, and T. Jebara
International Joint Conference on Artificial Intelligence (IJCAI), 2015

2014 "Collaborative Ranking for Local Preferences"
B. Kapicioglu, **D. Rosenberg**, R. Schapire, and T. Jebara
Proceedings of Artificial Intelligence and Statistics (AISTATS), 2014

2009 "Multi-View Point Cloud Kernels for Semi-Supervised Learning"
D. Rosenberg, V. Sindhwani, P. Bartlett, and P. Niyogi
IEEE Signals Processing Magazine, vol 26, no 5, pp 145-150, Sept 2009

2008 "An RKHS for Multi-View Learning and Manifold Co-Regularization"
V. Sindhwani and **D. Rosenberg**
International Conference on Machine Learning (ICML), 2008

2007 "Mixture-of-Parents Maximum Entropy Markov Models"
D. Rosenberg, D. Klein, and B. Taskar
Proceedings of Uncertainty in Artificial Intelligence (UAI), 2007

2007 "Rademacher Complexity of Co-Regularized Kernel Classes"
D. Rosenberg and P. Bartlett
Proceedings of Artificial Intelligence and Statistics (AISTATS), 2007

INVITED TALKS

2019	New England Statistics Symposium, May 16, 2019, Hartford, CT "Machine Learning for Structured and Unstructured Data in Finance" (*Keynote, joint with Amanda Stent*)
2019	Duke University Machine Learning Seminar Series, Apr 3, 2019, Durham, NC "Extracting Data from Tables and Charts in Natural Document Formats"
2018	International Chinese Statistical Association, Jun 14, 2018, New Brunswick, NJ "Extracting Data from Tables and Charts in Natural Document Formats"
2017	IEEE Computer Society, Rock Stars of ML and Deep Learning, Sep 12, 2017, Santa Clara, CA "Extracting Data from Tables and Charts in Natural Document Formats"
2012	NYAS Machine Learning Symposium, Oct 19, 2012, New York, NY "Location Challenge: Counting Visits to Starbucks"

OTHER PRESENTATIONS

2022	GPU Technology Conference (GTC), Mar 21, 2022 (Planned), Virtual "How Transformers Can Create Millions of Rows of Representative Synthetic Tabular Data"
2018	GPU Technology Conference (GTC), Mar 26, 2018, San Jose, CA "Extracting Data from Tables and Charts in Natural Document Formats"
2015	IAB Mobile Marketplace, Mar 30, 2015, New York, NY "Mobile Targeting Slam Dunk: Relevant Ads in Real-Time through Intent-Based Profiles"
2014	Mobile Media Summit, July 23-24, 2014, Chicago, IL "Mobile Retargeting, Optimization and Hitting the ROI Bullseye"
2014	Mobile Marketing Association, Mar 30, 2014, New York, NY "Mobile Targeting Slam Dunk: Relevant Ads in Real-Time through Intent-Based Profiles"
2013	Location Intelligence Summit, Mar 21-22, 2013, New York, NY "Mobile Location Data Quality"
2012	SAMSI Computational Advertising Workshop, Aug 6-17, 2012, Research Triangle Park, NC "Leveraging Location for Mobile Ad Targeting"

AWARDS

2019	**Neural Information Processing Systems, "Best Reviewer**" *given to top ~8% of reviewers*
2015 and 2016	**NYU Center for Data Science's Professor of the Year Award** *for "demonstrating a sincere desire to see the students succeed and committing to helping them achieve their goals, as voted on by the students."*
2001	**The MITRE Corporation Special Recognition Award** *for "the discovery of a key security vulnerability in a proposed ... design for a future generation of the Global Positioning System (GPS)"*
1998 and 2000	**Yale University's Anthony D. Stanley Prize** *for "excellence in pure and applied mathematics"*

1997 ***Tau Beta Pi* Engineering Honor Society**
for "top 1/8th of Technical Majors" at Yale College.

1994 **Eagle Scout**

REFEREE SERVICE

Association for the Advancement of Artificial Intelligence (AAAI), 2020
Conference on Computational Learning Theory (COLT), 2005, 2011
Conference on Uncertainty in Artificial Intelligence (UAI), 2009
IEEE Transactions on Information Theory, 2006, 2007, 2008
IEEE Transactions on Neural Networks, 2009, 2010
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2011
International Conference on Learning Representations (ICLR), 2019, 2020
Journal of Machine Learning Research (JMLR), 2005
Machine Learning (Springer), 2009, 2011
Neural Information Processing Systems (NeurIPS), 2008, 2009, 2011, 2014, 2016, 2017, 2018, 2019 ("**Best Reviewer**" – top ~8%), 2020

TEACHING EXPERIENCE

Fall 2021	New York University DS-GA 3001: Tools and Techniques for Machine Learning
Spring 2021	New York University DS-GA 3001: Tools and Techniques for Machine Learning
Spring 2019	New York University DS-GA 1003: Machine Learning
Spring 2018	New York University DS-GA 1003 / CSCI-GA 2567: Machine Learning
Fall 2017	Bloomberg FOML: Foundations of Machine Learning
Spring 2017	New York University DS-GA 1003: Machine Learning and Computational Statistics
Spring 2015	New York University DS-GA 1003: Machine Learning and Computational Statistics
Spring 2014	New York University DS-GA 1003: Machine Learning and Computational Statistics, project advisor for class taught by Prof. David Sontag
Spring 2008	UC Berkeley CS 281B/Stat 241B: Statistical Learning Theory, TA for Prof. Peter Bartlett
Fall 2006	UC Berkeley Stat 198: Teaching Statistics, TA for Prof. Deborah Nolan
Spring 2006	UC Berkeley Stat 210B: Theoretical Statistics, TA for Prof. Michael Jordan
Spring 2004	UC Berkeley Stat 20: Introduction to Probability and Statistics, TA for Dr. Hank Ibser
Fall 2003	UC Berkeley CS 281A/Stat 241A: Statistical Learning Theory, TA for Prof. Peter Bartlett
Fall 2003	UC Berkeley Stat 2: Introductory Statistics, TA for Prof. John Rice