

Lisa D. Friedland, Ph.D.

Boston, MA ♦ (413) 575-4995

LfriedlCS@gmail.com ♦ linkedin.com/in/lisafriedland ♦ Lfriedl.github.io

Data scientist and applied researcher. Skilled at developing new tools and extracting insights from real data. Broad expertise in data mining, machine learning, artificial intelligence, network science, anomaly and fraud detection, text analysis (NLP, information retrieval), computational social science. Innovative; clear communicator; thrives on collaboration.

SKILLS

Experience developing predictive models and fluency with core computer science areas.

Machine learning: statistical models (density estimation, probabilistic graphical models), methods for graph-structured data, clustering, classification and regression (deep learning, random / boosted forests, SVMs, imbalanced classes). **Programming languages:** most recent work in Python (*scipy*, *numpy*, *igraph*), R (*data.table*), SQL (*BigQuery*). Experienced with Java (*Weka*), C/C++, Perl, Matlab, LISP; Unix/Linux, Jupyter, GCP.

Design and management of projects manipulating large datasets (individually and in small teams): defining project scope, building software infrastructure, selecting and implementing algorithms, collecting and preprocessing data (web scraping, cleaning, feature engineering), developing evaluation measures, analyzing experimental results, creating & testing hypotheses.

Technical communication to varied audiences. Primary author of papers in *Science*, *ICML*, *KDD*, *EMNLP*, *SDM*, *CIKM*. Have taught computer science and math to undergraduates and high schoolers. Experienced with cross-functional and interdisciplinary work.

PROFESSIONAL EXPERIENCE

Google Cambridge, MA
Software Engineer, Ads Safety 2022

- Addressed data quality issues (inconsistencies in human labels) affecting the machine learning classifiers that detect policy-violating ads.

Northeastern University, Network Science Institute Boston, MA
Associate Research Scientist (epidemic-modeling group) 2020

- Developed measures to quantify people's movement patterns during the COVID era, using nationwide location/mobility data. Built data pipeline delivering them to a live dashboard.

Postdoctoral Research Associate (computer science group) 2019

- Applied my novel bipartite-graph methods to data from self-propagating malware.

Postdoctoral Research Associate (computational social science group) 2016–2018

Visiting Ph.D. Student 2013–2016

- Led a data linkage project matching 290 million public Twitter accounts to registered U.S. voters, yielding 1.8 million high-precision matches. By virtue of its known population properties, this panel forms the basis for numerous ongoing projects.
- Quick pivot after 2016 election to “fake news,” measured on Twitter using this data. Findings published in *Science* to wide media attention. Project included a lasso classifier to detect

political tweets (weakly supervised, retrained daily), careful manual validation of websites and users, management of tweet and website data, graph analysis of fake news sites.

- Additional: derived an EM classifier to resolve conflicting annotations of political stance; used tweets to predict vote share in polls; analyzed employment histories of Congressional staffers.

University of Massachusetts Amherst, Department of Computer Science Amherst, MA
Graduate Research Assistant, Knowledge Discovery Lab 2002–2013

- Analyzed and modeled large network-structured data sets. Developed methods for new types of data and new tasks (e.g., link prediction). Emphasis on explaining algorithm behavior through close attention to data characteristics and baseline models.
- Published 9 papers (4 as 1st author); presented talks and posters at international conferences and workshops; collaborated with external institutions on funded projects.
- Selected projects: detecting coordinated “tribes” from job history data, as an indicator for fraud risk in the securities industry; detecting insider threats via anomalous computer usage; unsupervised scores of node similarity in bipartite graphs; a search engine for half-remembered jokes.

LinkedIn Mountain View, CA
Intern, Analytics Group Summer 2010

- Examined when and why people connect: created a personalized “Connection Timeline” visualizing user history; analyzed shared affiliations among connected pairs of users.

PricewaterhouseCoopers San Jose, CA
Intern, Center for Advanced Research Summer 2009

- Designed graphical model to detect prescription drug fraud from health insurance claims.

Whitehead Institute / MIT Center for Genome Research Cambridge, MA
Scientific Programmer 1998–2000

- Designed, built and supported software pipeline enabling a large-scale survey of polymorphisms (SNPs) in human genes.

EDUCATION

University of Massachusetts Amherst Amherst, MA

- Ph.D., 2016. Dissertation: “Detecting anomalously similar entities in unlabeled data.”
- M.S., 2006. Master’s project: “Predicting protein-protein interactions using high-throughput data and network structures.”

Harvard University Cambridge, MA
A.B. with honors, Computer Science, 1998.

ACTIVITIES AND INTERESTS

Teaching, mentoring. Taught high school geometry at an international school in Rome (2005). Directed high school interns (2018–2020). Outreach activities at UMass for women and incoming students (2002–2013). **Reviewer** for scientific publications such as *Nature Human Behavior*, *ICWSM*, *SDM*, *IEEE Transactions on Knowledge and Data Engineering* (2017–2022). **Hobbies:** lifelong violin player (in orchestras). Languages: French (semester in Paris, worked in Brussels); have studied (≥ 1 yr, for fun) beginner-level Italian, German, Dutch, Spanish, Latin.