

Timothy N. Rubin, PhD

Data Science Team Lead

(215) 990-4012

Tim.Rubin@gmail.com

[Personal Website](#) | [Github Profile](#) | [Google Scholar](#)

PROFESSIONAL EXPERIENCE

DATA SCIENCE TEAM LEAD: CHANGE HEALTHCARE

2018-PRESENT

- Managed cross-functional team of 5+ data scientists and ML engineers working on healthcare-related ML products
- Developed and deployed numerous NLP models for automatically classifying healthcare data
- Contributed to a framework for building, evaluating, and deploying ML pipelines
- Contributed to an active-learning solution for efficiently collecting and labeling text data

SENIOR DATA SCIENTIST: SURVEYMONKEY

2016-2018

- Developed all algorithms underlying SurveyMonkey's [SurveyMonkey Genius](#) platform—a user-facing ML product that provides customers with predictions about their survey (e.g., estimated completion times) as well as personalized recommendations for how to improve their survey. SurveyMonkey Genius has received [industry press](#) as part of SurveyMonkey's 2017 company rebrand, and had significant financial and brand impact.
- Developed a new SurveyMonkey use-case ontology and automated classification model. Applications for this product include sales assistance and driving a variety of personalization opportunities for users.

SENIOR RESEARCH SCIENTIST: INDIANA UNIVERSITY

2013 – 2016

- Led and collaborated on research projects leading to numerous publications in top-tier journals and conferences
- Developed and implemented novel algorithms for (a) automatically identifying functional brain regions using GC-LDA (b) empirically evaluating semantic models, and (c) improving prediction methods for Latent Dirichlet Allocation models
- Secured a \$65,000 grant for studying linguistic features related to schizophrenia

GRADUATE RESEARCH SCIENTIST: UNIVERSITY OF CALIFORNIA, IRVINE

2006 - 2012

- Developed and implemented novel probabilistic topic models that achieved state-of-the-art performance on multi-labeled document classification
- Developed and implemented novel collaborative filtering algorithms for product ratings

RELEVANT SKILLS

Programming languages + Software Tools:

- Python; SQL; AWS; Spark; Docker; Hive; Java (some)

Statistical analysis software:

- R; MATLAB; Amazon QuickSight; SPSS; Excel

Analytical Skills:

- Machine learning; Natural language processing; Experimental research and design, Probability theory and statistics

Communication and Team Management Experience:

- Experience managing a team of 5+ people; Public speaking; Presentation of applied and theoretical research to both technical and non-technical audiences; Teaching

EDUCATION

University of California, Irvine

Ph.D., M.A., Department of Cognitive Sciences

Irvine, CA

2012, 2009

Tufts University

B.S. Cognitive Science

Medford, MA

May 2004

SELECTED PUBLICATIONS

Papanikolaou, Y., **Rubin, T.N.**, Tsoumakas, G. (2017) [Dense Distributions from Sparse Samples: Improved Gibbs Sampling Parameter Estimators for LDA](#). *Journal of Machine Learning Research (JMLR)*.

Rubin, T.N., Koyejo, O., Jones, M.N., Yarkoni, Y., (2016). [Generalized Correspondence-LDA Models \(GC-LDA\) for Identifying Functional Regions in the Brain](#). *30th Annual Conference on Neural Information Processing Systems (NIPS)*.

Rubin, T.N., Kievit-Kylar, B., Willits, J.A., Jones, M.N., (2014). [Organizing the Space and Behavior of Semantic Models](#), *36th Annual Conference of the Cognitive Science Society*.

Rubin, T.N., Chambers, A., Smyth, P., Steyvers, M., (2012). [Statistical Topic Models for Multi-Label Document Classification](#), *Machine Learning: special issue on Learning from Multi-Label Data*.

Rubin, T.N., Steyvers, M., (2009). [A Topic Model For Movie Choices and Ratings](#), *9th International Conference on Cognitive Modeling (ICCM)*, ([Supplementary Material](#))