Timothy N. Rubin, PhD

Data Science Leader
(215) 990-4012
Tim.Rubin@gmail.com
Personal Website | Google Scholar

SUMMARY

Accomplished data science leader with a PhD and over a decade of experience in machine learning, data analysis, and research. Proven track record in developing high-impact ML products, managing cross-functional technical teams, and driving data-informed decision-making in organizations. Committed to fostering an inclusive and collaborative work environment that encourages questions and diverse perspectives. Passionate about mentorship and empowering individuals for career growth.

PROFESSIONAL EXPERIENCE

VP OF DATA SCIENCE: DOMA 2023-PRESENT

DIRECTOR / SR. MANAGER OF DATA SCIENCE: DOMA 2020-2022

- Lead the multidisciplinary data science team at Doma consisting of ML engineers, data scientists and data analysts.
- Works closely with Doma's executive team to create innovative products in partnership with external companies and GSE's.
- Developer of multiple ML products, including the instant underwriting model which underwrites ~80% of Doma's $20-$30M annual enterprise revenue
- Managed relationships with executive and technical stakeholders in partner companies

DATA SCIENCE TEAM LEAD: CHANGE HEALTHCARE 2018-2020

- Led a cross-functional team of 5+ data scientists and ML engineers, with duties including technical leadership, project selection and management.
- Developed numerous healthcare-related ML products, bolstering Change's leadership in healthcare technology
- Developed DS project lifecycle standards that were adopted throughout the AI group

SENIOR DATA SCIENTIST: SURVEYMONKEY (NOW MOMENTIVE.AI) 2016-2018

- Developed the algorithmic backbone underlying SurveyMonkey's SurveyMonkey Genius—a customer-facing ML product providing predictive analytics and personalized recommendations for surveys.
- Created an internal use-case ontology and automated classification model, driving sales assistance and user personalization opportunities.

SENIOR RESEARCH SCIENTIST: INDIANA UNIVERSITY 2013 – 2016

- Led and collaborated on research projects leading to numerous publications in top-tier journals and conferences. Notable work includes: (a) creating the GC-LDA model and applying it for automated identification of functional brain regions (b) empirical evaluations and comparisons of semantic models, and (c) improving prediction methods for Latent Dirichlet Allocation models.

GRADUATE RESEARCH SCIENTIST: UNIVERSITY OF CALIFORNIA, IRVINE 2006 – 2012

- Developed, implemented, and published probabilistic machine learning models for multi-label document classification, recommendation systems, and leveraging metadata plus text for concept abstractions.

RELEVANT SKILLS

Programming: Python; SQL; AWS; Spark; Docker; Java (some)

Statistical Analysis Tools: Python scientific stack; R; MATLAB; Excel; SPSS

Analytical Skills: Machine learning; Research; Experimental Design; Probability theory and statistics; NLP

Communication and Leadership Experience: Experience managing teams of 6+ people including ML engineers, data scientists, and data analyses; Extensive experience working with and presenting technical materials to executive teams and technical audiences

EDUCATION

University of California, Irvine — Irvine, CA
Ph.D., M.A., Department of Cognitive Sciences — 2012, 2009

Tufts University — Medford, MA
B.S. Cognitive Science — May 2004

SELECTED PUBLICATIONS & PATENTS (FULL LIST ON PERSONAL WEBSITE)

Mason, E.K., **Rubin, T.N.** (2022) Predictive time series data object machine learning system. United States Patent US US11580309B1. United States Patent and Trademark Office.

Papanikolaou, Y., Foulds, J. R., **Rubin, T. N.**, & Tsoumakas, G. (2017). Dense distributions from sparse samples: improved Gibbs sampling parameter estimators for LDA. *The Journal of Machine Learning Research*, 18(1), 2058-2115.

Rubin, T. N., Koyejo, O. O., Jones, M. N., & Yarkoni, T. (2016). Generalized correspondence-LDA models (GC-LDA) for identifying functional regions in the brain. *Advances in neural information processing systems*, 29.

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. Cognitive Science Society (US). Conference (Vol. 2014, p. 1329). NIH Public Access.

Rubin, T. N., Chambers, A., Smyth, P., & Steyvers, M. (2012). Statistical topic models for multi-label document classification. *Machine learning*, 88, 157-208.

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