# Timothy N. Rubin, PhD

Machine Learning & Data Science Leader

(215) 990-4012

Tim.Rubin@gmail.com

Personal Website | Google Scholar


## SUMMARY


Accomplished data science leader with over nearly two decades of experience in AI, machine learning, data analytics and research. Proven track record in building high-impact machine learning products, growing and managing cross-functional teams, and driving AI strategy and data-based decision making throughout large organizations. Extensive experience working with executive teams and collaborating with both technical and non-technical stakeholders. Passionate about mentorship, and empowering individuals for career growth.


## PROFESSIONAL EXPERIENCE

### FOUNDER & CEO, WELLNESS TECHNOLOGY INC. 2024-PRESENT

- As a solo founder, built an AI-powered mental health app Wellness AI for Android and iOS, with thousands of monthly active users, $5k MRR and a 4.9/5 rating worldwide across over 200 user reviews
- Lead product, engineering, and marketing efforts, while managing engineering and marketing contractors

### AI CONSULTANT 2023-2025

- Partnering with a diverse portfolio of clients across different growth phases, from pre-seed startups to established public companies, to formulate and implement generative AI strategies.
- Assisted organizations in constructing actionable roadmaps for the integration of generative AI into production systems, focusing on immediate benefits and long-term customer value.
- Developed prototypes of systems including RAG-based LLM and traditional classifiers

### SR. MANAGER / DIRECTOR / VP OF DATA SCIENCE: DOMA 2020-2023

- Led cross functional data science team consisting of ML engineers, data scientists and data analysts.
- Collaborated closely with Doma's executive team to develop innovative machine learning products and business initiatives.
- Designed, built and deployed multiple ML products, including the risk model which underwrote ~80% of Doma's $20-$30M annual enterprise revenue, and deep learning models to support operational efficiency.
- Managed partnership with technical teams from Fannie Mae and federal agencies to get regulatory approval for Doma's Title Waiver program.

### DATA SCIENCE TEAM LEAD: CHANGE HEALTHCARE 2018-2020

- Led a cross-functional team of ML engineers and data scientists, with duties including technical leadership, project management, and team 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 2016-2018

- Developed the algorithmic backbone underlying the SurveyMonkey Genius platform—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 2012 – 2016

- Led and collaborated on research projects leading to numerous publications in top-tier journals and conferences in machine learning and other fields. Notable work includes: (a) developing a novel topic model, GC-LDA, model and leveraging it to jointly analyze text and spatial data from neuroscience publications (b) empirical evaluations and comparisons of semantic models, and (c) improving Bayesian 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

**Communication and Leadership Experience:** Five years of experience managing teams consisting of ML engineers, data scientists and data analysts, including other managers. Extensive experience working directly with executive leadership and presenting technical material to both non-technical and technical audiences.

**Areas of Expertise:** Machine learning; Analytics; LLMs; Natural Language Processing; Prompt Engineering; Research; Experimental Design; Probability theory and Statistics

**Programming and Analysis Tools:** Python; SQL; Spark; Python scientific stack; PyTorch; R; MATLAB; Excel; SPSS; Looker

## 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)