

TRANG NGUYEN

<https://www.linkedin.com/in/trangnguyen17/>

EDUCATION

University of Massachusetts, M.S./Ph.D. Computer Science in progress	Expected 2028
<ul style="list-style-type: none">• Advisor: Prof. Madalina Fiterau• GPA: 4.0 / 4.0	
Grinnell College, B.A. Mathematics & Statistics with Honors (GPA: 3.94/4.00)	2017

HONORS AND AWARDS

- Honorable Mention for Ford Foundation Predoctoral Fellowship 2023
- Lori A. Clarke Scholarship 2023
- Finalist in the Three Minute Thesis Competition 2023 2023
- Phi Beta Kappa, Academic Honor Society 2017
- First Place, National Undergraduate Class Project Competition (American Statistics Association) 2016
- Summer Business Scholars Program Full Scholarship, University of Chicago 2015

KNOWLEDGE AND SKILLS

- Knowledge : Machine Learning, Network Analysis, Survey Methodology, Experimental Design
- Programming : Python, Java, R, SQL (*Advanced*), HTML & CSS, Javascript (*Intermediate*), C (*Basic*)
- Languages : English (fluent), Vietnamese (fluent), Korean (basic), Japanese (basic)
- Others : Software development, Database, Spark, Cloud Computing

SELECTED PUBLICATIONS

- Vela D., Sharp A., Zhang R., **Nguyen T.**, Hoang A., Pianykh O.S.. Temporal quality degradation in AI models. *Sci Rep* **12**, 11654 (2022).
- Bailey P., Lee M., **Nguyen T.**, Zhang T. (2020) Using EdSurvey to Analyse PIAAC Data. In: Maehler D., Rammstedt B. (eds) Large-Scale Cognitive Assessment. Methodology of Educational Measurement and Assessment. Springer, Cham.
- Long, Y.*, **Nguyen, T.***, & Tareque, I.* (2016). Logistic Regression and Classification Tree on Customer Churn in Telecommunication. First Place Award, National Undergraduate Statistics Class Project Competition. (* equally contributed)

SELECTED ONGOING AND UNPUBLISHED RESEARCH PROJECTS

- Analyzing Eating Disorders through Social Media Anecdotes, UMass Amherst** 2023 - Present
- Analyzed different triggers and recovery patterns of people who suffer from eating disorders and self-report on Reddit from 2020 to 2023 using topic modeling and large language models (GPT-3)
 - Abstract submitted to TADA 2023
- Optimal Genomics-Drug Fusion in Predicting Cancer Drug Response, UMass Amherst** 2022 - 2023
- Analyzed different fusion methods to combine genomics and drug embeddings and tested how different methods react to different fusion points
 - Manuscript to be submitted to Briefings in Bioinformatics
- Informal Network Analysis of Staff in a College, Grinnell College** 2016
- Designed survey and used network analysis to design a cost-effective method to group staff in an organization.
 - Delivered a white paper and web application to the Human Resources Department of Grinnell College

INDUSTRY EXPERIENCE

Lead DataOps Engineer, Tamr Inc. 2019 – 2022

- Worked with business executives and software architects to design, implement, and productionize data pipelines, integrating the machine learning solution in the data architecture ecosystem.
- Trained semi-supervised clustering models with active learning by engaging subject-matter experts in an iterative process of training and validating machine learning models.
- Designed and programmed microservices to run regression testing on big-scale clinical mapping projects, and web apps to monitor model drift

Data Scientist Assistant, American Institutes for Research, *Washington, D.C.* 2017 – 2019

- Co-authored two published R packages ([EdSurvey](#) and [WeMix](#)) that connects to big survey databases and analyzes complex survey data with imputation variance and multilevel weights.
- Implemented deep learning algorithms to extract behavioral features from videos, and natural language processing to extract features from video transcripts for classroom analytics.

SERVICE & OUTREACH

- Advisor for Voices of Data Science 2023 - Present
- Co-Chair for Voices of Data Science Conference 2022 – 2023
- Data Science Talent Fellow at Open Avenues Foundation 2020 - 2021