

# Forough Poursabzi-Sangdeh

forough.poursabzi@microsoft.com • <https://foroughp.github.io>

Microsoft Research

641 Avenue of the Americas, 7th Floor

New York, NY 10011

## RESEARCH INTERESTS

---

I am interested in the interaction between humans and machine learning models. Because of the heavy involvement of humans throughout the machine learning lifecycle—from getting data that is necessary to train models to making decisions with the help of models—it is becoming more and more important to study humans and machines jointly. My research builds on research from HCI, psychology, and social sciences on human-subject experiments to study human behavior when they interact with machine learning models. Recently, I have focused on studying this interaction in the context of interpretability and fairness. My long-term ambition is to leverage insights from these studies for creating systems that foster an effective and responsible collaboration between humans and models.

## EDUCATION

---

**University of Colorado**, Boulder, CO

PhD in Computer Science, 2013–2018

Advisor: Jordan Boyd-Graber

Dissertation: Design and Empirical Evaluation of Interactive and Interpretable Machine Learning

Master of Science, Computer Science, 2013–2015

**University of Tehran**, Tehran, Iran

BE in Computer Engineering, 2008–2012

## EMPLOYMENT

---

**Microsoft Research**, New York, NY

Postdoctoral Researcher, July 2018–Present

**Microsoft Research**, New York, NY

Research Intern, June 2017–August 2017

Supervisor: Jennifer Wortman Vaughan

Research on studying model interpretability from human perspective using human-subject experiments

**Oracle Labs**, Burlington, MA

Research Intern, June 2016–August 2016

Supervisor: Pallika Kanani

Research on text classification with minimally labeled documents using active learning and topic models

**University of Colorado**, Boulder, CO

Graduate Research Assistant, August 2014–July 2018

Supervisor: Jordan Boyd-Graber

Research on designing, implementing, and evaluating systems to help humans label and understand large document collections

## PUBLICATIONS

---

**Forough Poursabzi-Sangdeh**, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wortman Vaughan, and Hanna Wallach. “Manipulating and Measuring Model Interpretability”. *Revise and resubmit at Management Science*. (A shorter version appeared at the NIPS workshop on Transparent and Interpretable Machine Learning in Safety Critical Environments, 2017.)

Bran Knowles, Alison Smith-Renner, **Forough Poursabzi-Sangdeh**, Di Lu, Halimat Alabi. “Attending to the Problem of Uncertainty in Current and Future Health Wearables”. *Communications of the ACM (CACM)*, 2018.

Alison Smith, Tak Yeon Lee, **Forough Poursabzi-Sangdeh**, Leah Findlater, Jordan Boyd-Graber, and Niklas Elmqvist. “Evaluating Visual Representations for Topic Understanding and Their Effects on Manually Generated Labels”. *Transactions of the Association for Computational Linguistics (TACL)*, 2017.

**Forough Poursabzi-Sangdeh**, Jordan Boyd-Graber, Leah Findlater, Kevin Seppi. “ALTO: Active Learning with Topic Overviews for Speeding Label Induction and Document Labeling”. *Association for Computational Linguistics (ACL)*, 2016. [Code, now being used by Snagajob]

Alison Smith, Tak Yeon Lee, **Forough Poursabzi-Sangdeh**, Jordan Boyd-Graber, Niklas Elmqvist, Kevin Seppi, Leah Findlater. “Human-Centered and Interactive: Expanding the Impact of Topic Models”. *HCML Workshop at CHI*, 2016.

Jason Chuang, John D. Wilkerson, Rebecca Weiss, Dustin Tingley, Brandon M. Stewart, Margaret E. Roberts, **Forough Poursabzi-Sangdeh**, Justin Grimmer, Leah Findlater, Jordan Boyd-Graber, and Jeffrey Heer. “Computer-Assisted Content Analysis: Topic Models for Exploring Multiple Subjective Interpretations”. *NIPS Workshop on Human-Propelled Machine Learning*, 2014.

**Forough Poursabzi-Sangdeh**, and Ananth Kalyanaraman. “On Clustering Heterogeneous Networks”. *SIAM Workshop on Network Science*, 2013.

## INVITED TALKS

---

### **Manipulating and Measuring Model Interpretability**

Women in Machine Learning and Data Science (WiMLDS) Meetup, New York, May 2019

O’Reilly Artificial Intelligence Conference, New York, NY, April 2019

Machine Learning Conference (MLConf), San Francisco, CA, November 2018

### **Design and Empirical Evaluation of Interactive and Interpretable Machine Learning**

NYU Tandon School of Engineering, New York, NY, April 2019

Department of Information Science, Boulder, CO, March 2018

Microsoft Research, New York, NY, March 2018

Oracle Labs, Burlington, MA, March 2018

## PROFESSIONAL SERVICE

---

### **Reviewer and Committees**

AAAI 2020, IJCAI 2019 and 2020, CSCW 2019, NeurIPS 2019 (selected as a top reviewer), NeurIPS 2019 Reproducibility Challenge, ICML 2019, ACM Transactions on Computing Education 2018, AIS-TATS 2018, EMNLP 2018, COLING 2018, EACL 2016 and 2017, Debugging Machine Learning Models workshop at ICLR 2019, Explainable Smart Systems workshop at Intelligent User Interfaces (IUI) 2018, 2019, 2020, AI for Social Good workshop at NeurIPS 2019, Advances in Interpretable Machine Learning and Artificial Intelligence workshop at KDD 2019, Black in AI workshop at NeurIPS 2017 and 2018

## **Workshop Organizer**

Senior Program and Mentorship Chair for the Women in Machine Learning (WiML) workshop co-located with NeurIPS 2019

## **MENTORING**

---

**Microsoft Research Data Science Summer School Program**, 2019

Research mentor for a group of eight NYC undergrad students for a research project on replication

**Harpreet Gaur**, undergraduate student at CUNY City Tech, 2019

**You Lu**, masters student at University of Colorado, 2016 (current position: PhD student at Virginia Tech)

## **MISCELLANEOUS**

---

### **Speaker**

“It’s time for data scientists to collaborate with researchers in other disciplines”, The O’Reilly Data Show Podcast, 2019.