A N N A   F A R I H A

Homepage: people.cs.umass.edu/afariha/
Email: afariha@cs.umass.edu

RESEARCH INTEREST

The aim of my research is to democratize data and data systems towards boosting productivity and agility. To this end, my research focuses on three key aspects of data system democratization: enhancing *usability* of data systems for non-experts, providing *explanation* frameworks to enable understanding of certain outcomes, and achieving *trust* and *fairness* in data-driven machine learning.

- **Data systems usability:** query by example, sampling by example, document summarization by example.
- **Debugging:** root cause analysis of nondeterministic failure, interventional debugging, data debugging.
- **Causal reasoning and explanation:** explaining change in data, explaining data non-conformance.
- **Trust and fairness:** trusted machine learning, fair machine learning.
- **Metadata management and its applications:** data constraints, data drift.

EDUCATION

- Doctor of Philosophy in Computer Science (in progress)
  College of Information and Computer Sciences, University of Massachusetts, Amherst
  Dissertation title: Enhancing Usability and Explainability of Data Systems
    - Advisor: Alexandra Meliou
    - Committee members: Peter Haas (UMass), Emery Berger (UMass), Suman Nath (MSR)
    - Expected graduation: May 2021
- Master of Science in Computer Science (2020)
  College of Information and Computer Sciences, University of Massachusetts, Amherst
    - CGPA: 4.00 (in the scale of 4.00)
- Master of Science in Computer Science and Engineering (2011)
  University of Dhaka, Bangladesh
    - CGPA : 4.00 (in the scale of 4.00)
- Bachelor of Science in Computer Science and Engineering (2010)
  University of Dhaka, Bangladesh
    - CGPA: 3.83 (in the scale of 4.00)

EMPLOYMENT HISTORY

- **Research Assistant** (September 2016 – to date)
  Data Systems Research for Exploration, Analytics, and Modeling (DREAM) Lab, CICS, UMass
- **Research Intern** (Summer 2019 & 2018)
    - PROSE Research & Engineering, Microsoft (Mentor: Ashish Tiwari)
    - DMX Lab, Microsoft Research, Redmond (Mentor: Suman Nath)
- **Lecturer** (2012 – 2016)
    - Computer Science and Engineering, University of Dhaka, Bangladesh (2014 – 2016)
    - Computer Science and Engineering, United International University, Bangladesh (2012 – 2014)
- **Software Engineer** (2011 – 2012)
  Structured Data Systems Limited, Dhaka

HONORS, AWARDS, & GRANTS

- VLDB Best Demonstration Runner-up award, 2020
- Microsoft Research Dissertation Grant, 2020
- Travel and accommodation award in CRA-W Grad Cohort Workshop, Washington DC, 2017
- Jim Gray scholarship, CICS, UMass Amherst, 2016
- Special Achievement Award – Bangladesh ICT Journalist Forum, 2012
- Special Achievement Award – Bangladesh Women in IT, 2012
- Special Achievement Award for outstanding performances in programming contests, 2011
- Dean's Honor Award and Dean's Award – Faculty of Engineering, University of Dhaka (top 1%), 2010
- Fatema-Iqbal trust scholarship, University of Dhaka, 2008
- National Education Board Scholarship: 12th grade (2006) & 10th grade (2004)
- Best Student of the School (Kushtia Govt. Girls' High School), 2003
- Talent Pool Scholarship in National Scholarship Examination: 8th grade (2001) & 5th grade (1998)

PUBLICATIONS

1. A Fariha, A Tiwari, A Radhakrishna, S Gulwani, A Meliou. Conformance Constraint Discovery: Measuring Trust in Data-Driven Systems. (under review).
2. A Fariha, L Cousins, N Mahyar, A Meliou. Example-Driven User Intent Discovery: Empowering Users to Cross the SQL Barrier Through Query by Example. (under review).
3. MT Islam, A Fariha, A Meliou. Through the Data Management Lens: Experimental Analysis and Evaluation of Fair Classification. VLDB 2021 (under review).
4. A Fariha, S Nath, A Meliou. Causality-guided Adaptive Interventional Debugging. SIGMOD 2020.
5. A Fariha, M Brucato, PJ Haas, A Meliou. SuDocu: Summarizing Documents by Example. VLDB 2020 (Best demonstration runner-up award).
6. A Fariha, A Tiwari, A Radhakrishna, S Gulwani. ExTuNe: Explaining Tuple Non-conformance. SIGMOD 2020.
7. A. Fariha, A. Meliou. Example-Driven Query Intent Discovery: Abductive Reasoning using Semantic Similarity. VLDB 2019.
8. A Fariha, SM Sarwar, A Meliou. SQuID: Semantic Similarity-Aware Query Intent Discovery. SIGMOD 2018.
9. A Fariha, CF Ahmed, CK Leung, M Samiullah, S Pervin, L Cao. A New Framework for Mining Frequent Interaction Patterns from Meeting Databases. Engineering Applications of Artificial Intelligence 2015.
10. M Samiullah, CF Ahmed, A Fariha, A Uddin: Efficient Graph Classification in Shifted Datasets Using Weighted Correlated Feature Selection. LMCE Workshop, ECML PKDD 2014.
11. A Fariha, CF Ahmed, CK Leung, SM Abdullah, L Cao: Mining Frequent Patterns from Human Interactions in Meetings Using Directed Acyclic Graphs. PAKDD 2013.

SERVICE & OUTREACH

- **Program committee member:** SIGMOD 2021 (research track), EDBT 2021 (demo track)
- **External reviewer**: SIGMOD 2018, 2019, and 2020; CHI 2021
- **Mentor**: Research Experience for Undergraduate Students (REU), UMass, Summer 2017
- **Volunteer**: Girls Inc. Eureka! Workshop, Summer 2017

OTHER PROJECTS

- Data change summarization (ongoing)
- Exposing the disconnect between data and data-driven systems (ongoing)
- Example-driven data sampling (ongoing)
- Maximum coverage problem with path constraint (Fall 2018)
- Automatic image captioning using multi-task learning (Fall 2017)
- Neural model for contextual named entity retrieval (Fall 2017)
- Human activity recognition from sensor data (Spring 2017)
- Accelerating query evaluation using multi-query prediction (Fall 2016)

GRADUATE-LEVEL COURSES

Advanced Algorithms, Database Design and Implementation, Machine Learning, Neural Networks: A Modern Introduction, Algorithms for Data Science, Probabilistic Graphical Model, Parallel Algorithms and Architecture, Approximation Algorithms and Combinatorial Optimization.

COMPETITIVE PROGRAMMING & PROBLEM SOLVING

Event	Position	Team	Year
ACM ICPC (Dhaka Regional, Bangladesh)	5	DU Resonance	2011
IUT ICT Fest Programming Contest (Dhaka, Bangladesh)	1	DU Resonance	2011
DU NCPC (Dhaka, Bangladesh)	1	DU Resonance	2011
DU NCPC – Women (Dhaka, Bangladesh)	1	DU Felicity	2011
ACM ICPC (Amritapuri Regional, India)	6	DU Resonance	2010
ACM ICPC (Dhaka Regional, Bangladesh)	4	DU Resonance	2010
ACM ICPC (Dhaka Regional, Bangladesh)	3	DU Ouranos	2009
Google Code Jam	515	Individual	2009

- **Problem Setter & Judge**
  - SAARC Programming Contest, 2014 & 2013, Bangladesh University
  - Bangladesh Informatics Olympiad, 2013, National Round
  - ACM ICPC – Dhaka Site, 2012, Daffodil International University

TECHNICAL SKILLS

- Python, LaTeX, Java, SQL, C, C++, PHP, Javascript, R