Manisha Mukherjee

Email: mmukherj@andrew.cmu.edu
School of Computer Science
Carnegie Mellon University

Interests

- Machine Learning applications in Software Engineering, Security, NLP.

Education

Carnegie Mellon University — Pittsburgh, PA
PhD in Software Engineering, GPA: 3.95/4.00 — 2017–Current
- Advisor: Dr. Vincent Hellendoorn

The Pennsylvania State University — University Park, PA
M.S. in Computer Science and Engineering, GPA: 3.72/4.00 — 2012–2014
- Advisor: Dr. Thomas La Porta
- Thesis: "Determination of real-time traffic flow parameters in different devices based on QoI requirements"

West Bengal University of Technology — Kolkata, India
B.Tech in Computer Science and Engineering, GPA: 8.19 — 2007–2011

Scholarships and Awards

- Presidential Fellowship in SCS — 2023
- Carnegie Institute of Technology Dean's Fellowship — 2017
- Frank J. Marshall Graduate Fellowship — 2018
- Center for Integrated Healthcare Delivery Systems (CIHDS) Scholarship — 2012

Experience

Carnegie Mellon University — Pittsburgh, PA
Graduate Research Assistant — August 2017-Current
- ML4SE: Developer forum analysis, mining software repositories, code recommendation and code security.

Lawrence Livermore National Laboratory — Livermore, CA
Research Intern — Summer 2023
- Deep learning for HPC error log classification

Lawrence Livermore National Laboratory — Livermore, CA
Research Intern — Summer 2022
- Deep learning for power net load forecasting

Fujitsu Labs America — Sunnyvale, CA
Research Intern — Summer 2021
- Empirical Study on Kaggle kernels.
- Mining data from Kaggle and analyzing popularity metrics for Kaggle kernels to guide the enhancement of pipelines produced by Fujitsu's AutoML core pipeline synthesis by identifying pipeline features highly desired by data scientists.

Idaho National Laboratory Idaho Falls, ID
Research Intern, INL Wireless Security Institute Summer 2020

- Wireless signal classification and threat detection. Designed and built a tool for spectrum monitoring and threat visualization in real-time.

Fujitsu Labs America Sunnyvale, CA
Research Intern Summer 2019

- Semantic code search and code recommendation using Deep Learning.

Cisco Systems, Inc San Jose, CA
Software Engineer, ASR9K group October 2014- August 2017

- Developed a distributed router with an intelligent control plane (ICON) by using SDN technologies to dis-aggregate control plane and data plane.
- Wrote ODL (OpenDayLight) plugins and REST APIs to control several underlying router clusters.

The Pennsylvania State University University Park, PA
Graduate Research Assistant (Networking and Security Research Center) August 2012-August 2014

- Worked on extracting features from traffic videos and determining real-time traffic flow parameters in different devices.
- Optimized the transfer and processing of this data based on Quality of Information (QoI) requirements.

Capgemini India Pvt. Ltd. Kolkata, India
Senior Software Engineer July 2011-July 2012

Teaching

- **Teaching Assistant** at Carnegie Mellon University Fall 2022
  Principles of Software Construction Objects, Design, and Concurrency (17-214)
- **Head Teaching Assistant** at Carnegie Mellon University Fall 2019
  INI MSIT Project Practicum (14-798)
- **Teaching Assistant** at Pennsylvania State University Fall 2013
  Communication Networks (CMPEN 362)

Publications

[1] **M. Mukherjee** and V. J. Hellendoorn, "Stack over-flowing with results: The case for domain-specific pre-training over one-size-fits-all models", *arXiv preprint arXiv:2306.03268*, 2023.

[2] V. J. Hellendoorn, J. Tsay, **M. Mukherjee**, and M. Hirzel, "Towards automating code review at scale", in *29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '21)*, 2021.

[3] **M. Mukherjee**, M. Bahrami, and W. P. Chen, "Source code retrieval", in *US Patent Application 17/085,894*, 2020.

[4] **M. Mukherjee**, J. Edwards, H. Kwon, and T. F. La Porta, "Quality of information-aware real-time traffic flow analysis and reporting", in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, IEEE, 2015, pp. 69–74.

[5] **M. Mukherjee**, "Determination of real-time traffic flow parameters in different devices based on qoi requirements", in *MS Thesis*, 2014.

[6] Y. Han, **M. Mukherjee**, K. Yi, N. Yao, and C. Tucker, "Predicting early stage disease progression in patients with neurological disorders using high dimensional sensor data", in *Technical Report, Center for Integrated Healthcare Delivery Systems (CIHDS) Workshop*, 2013.