Monoshiz Mahbub Khan

Personal website ◇ LinkedIn profile
monoshizmk@gmail.com

EDUCATION

PhD in Computing and Information Sciences
Rochester Institute of Technology
August 2021 - Present (Expected: May 2026)
CGPA: 3.93 (out of 4.00)

BSc in Computer Science and Engineering
University of Dhaka
January 2016 - January 2020
CGPA: 3.55 (out of 4.00)

PUBLICATIONS

- Khan, M. M., & Yu, Z. (2024). Approaching Code Search for Python as a Translation Retrieval Problem with Dual Encoders. *Empirical Software Engineering*, 30(1), 1-28. DOI: 10.1007/s10664-w024-10580-3
- Khan et al. Efficient Story Point Estimation With Comparative Learning. arXiv preprint arXiv:2507.14642 (2025).

WORK EXPERIENCE

- **Intern** **June 2024 - August 2024**
  ABB
  Mannheim, Germany
  – Developed an end-to-end Named Entity Recognition pipeline using a range of techniques, including traditional NLP methods, ML models, deep learning models, LLMs and generative AI tools
  – Used various tools including PyTorch, spaCy, scikit-learn, Hugging Face, MLflow
  – Conducted under the supervision of Nika Strem as part of the DAAD RISE Professional Program 2024

- **Graduate Research Assistant** **Fall 2021 - Fall 2023, Fall 2024, Fall 2025**
  Lab of Human-In-the-Loop Software Engineering
  Rochester Institute of Technology
  Supervisor: Dr. Zhe Yu
  – Conducted research on **code search** and **comparative learning**, using deep learning, NLP and ML methods
  – Also explored research topics involving LLMs, image processing, software engineering and generative AI
  – Mentored Masters students on thesis projects: guided experimental design, advised research direction, and provided feedback on thesis writing

- **Graduate Teaching Assistant**
  Rochester Institute of Technology
  – **IDAI-720: Research Methods for Artificial Intelligence** **Spring 2024**
    * Instructors: Dr. Zhe Yu & Dr. Esa Rantanen
    * Graded assignments and final projects, and hosted office hours
  – **IDAI-710: Fundamentals of Machine Learning** **Spring 2025**
    * Instructor: Dr. James Heard
    * Graded assignments, hosted office hours, conducting review classes

DOCTORAL RESEARCH EXPERIENCE

Code Search 2021 - 2024

Research project focusing on retrieving programming language artifacts related to some natural language queries from a pool of possible programming language artifacts, using dual encoder models. Model is built in Python using TensorFlow and Keras modules. This work has been published in EMSE and presented at FSE 2025 in the journal-first track.

Comparative learning 2023 - Present

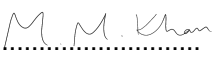
Research project focusing on modeling learning comparative judgments for Agile story point estimation. Experiments involved using GPT2, SBERT, and other deep learning and large language model (LLM) based structures and traditional machine learning methods. The framework was built using TensorFlow modules. The research has been conducted under the guidance of Dr. Zhe Yu. This work was recently submitted for peer-review at ICSE 2026.

TECHNICAL STRENGTHS

Programming languages	Python, JAVA, R, C, C++, JavaScript
Machine Learning & AI	TensorFlow, Keras, PyTorch, scikit-learn
MLOps & Data Engineering	MLflow, Airflow, PySpark
Frameworks & Databases	Flask, Spring, Angular, SQL (Oracle, SQLite), NoSQL (mongoDB)
Tools & Methodologies	Git (GitHub, Azure DevOps), LaTeX, Agile, Scrum

SCHOLARSHIPS

- DAAD Rise Professional Program 2024



Signature