Atharva Naik

arnaik@andrew.cmu.edu atharva-naik Atharva Naik
https://atharva-naik.github.io/ Google Scholar

Education

2024 – · · · · **Ph.D. Language Technologies, Carnegie Mellon University**.
Advisors: Carolyn Rose, Daniel Fried

2022 – 2024 **M.S. Language Technologies, Carnegie Mellon University**
GPA: 4.08/4

2018 – 2022 **B.Tech. Computer Science, Indian Institute of Technology, Kharagpur**
GPA: 9.66/10

Research

Statement: I want to build **Large Language Model (LLM) agents** that can **collaborate** with humans for software engineering while facilitating **on-the-job learning** about code efficiency, maintainability, and security. I want to help programming students and software engineers to be more productive and become better programmers at the same time.

Interests: Natural Language Processing (NLP), Deep Learning for Code (DL4C), Large Language Models of Code, LLM Agents, Interactive and Repository Level Code Generation, Editing, Translation and Review.

Experience

2022 – 2024 **Research Assistant, Carnegie Mellon University** Automated and Interactive Code Intelligence — Code Generation, Translation, and Review for assisting Software Engineers and Programming Students.

2021 **Research Intern, Technische Universität Darmstadt** Neural Network Arcthitecture for Faithful Interpretability in NLP.

Research Intern, Adobe RL agent for Creative Human-Human Collaboration.

Research Intern, University of Alberta Neuro-Symbolic Fuzzy Logic-based Reasoning for Explainable Natural Language Inference.

2019-2020 **Student Researcher, Autonomous Ground Vehicle (AGV) Group** Path Planning and Localization for Autonomous Driving.

Publications

Conference Publications

1. **A. Naik**, J. R. Yin, A. Kamath, *et al.*, "Generating situated reflection triggers about alternative solution paths: A case study of generative ai for computer-supported collaborative learning," in ***AIED***, 2024.
2. A. Rao, S. Vashistha, **A. Naik**, S. Aditya, and M. Choudhury, "Tricking LLMs into Disobedience: Understanding, Analyzing, and Preventing Jailbreaks," in ***LREC-COLING***, 2024.
3. **A. Naik**, S. Das, J. Vedurada, and S. Aditya, "Sync: A structurally guided hard negative curriculum for generalizable neural code search," in ***AACL***, 2023.
4. Z. Wu, Z. X. Zhang, **A. Naik**, Z. Mei, M. Firdaus, and L. Mou, "Weakly Supervised Explainable Phrasal Reasoning with Neural Fuzzy Logic," in ***ICLR***, 2023.
5. Y. Xie, **A. Naik**, D. Fried, and C. Rose, "CMTrans: Improving Code Translation with Comparable Corpora and Multiple References," in ***EMNLP Findings***, 2023.

6. S. Bv, J. A. Patel, **A. Naik**, Y. Butala, S. Sharma, and N. Chhaya, "Towards Enabling Synchronous Digital Creative Collaboration: Codifying Conflicts in Co-Coloring," in ***CHI Extended Abstracts***, 2022.
7. B. Santra, S. Roychowdhury, A. Mandal, *et al.*, "Representation Learning for Conversational Data using Discourse Mutual Information Maximization," in ***NAACL***, 2022.
8. Y. Wang, S. Mishra, P. Alipoormolabashi, *et al.*, "Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ NLP Tasks," in ***EMNLP***, 2022.
9. R. Mukherjee, **A. Naik**, S. Poddar, S. Dasgupta, and N. Ganguly, "Understanding the Role of Affect Dimensions in Detecting Emotions from Tweets: A Multi-task Approach," in ***SIGIR***, 2021.

Preprints

1. **A. Naik**, M. Alenius, D. Fried, and C. Rose, "Crscore: Grounding automated evaluation of code review comments in code claims and smells," 2024. arXiv: 2409.19801 [cs.SE]. URL: https://arxiv.org/abs/2409.19801.
2. **A. Naik**, K. Zhang, N. Robinson, *et al.*, "Can large language models code like a linguist?: A case study in low resource sound law induction," 2024. arXiv: 2406.12725 [cs.CL].

Under Review

1. **A. Naik**, J. R. Yin, A. Kamath, *et al.*, *Generative ai enabled situated reflection triggers for catalyzing solution path exploration*, 2024.

Projects

Best Practice Recommendation Agent for Python Code (Ongoing)

- Developing an agent for the recently introduced task of best practice recommendation for code review conditioned on Python version to minimize outdated best practices.
- Formulating the task as knowledge conflict mitigation to detect best practice violations for various Python versions faithfully.
- Creation of a benchmark and parameter-efficient inference and fine-tuning methods for best practice recommendation with faithful attribution of knowledge sources. (Planned)

Reference Free Evaluation of Automated Code Review Comment Generation

- We create CRScore, the first reference-free automated metric for code review comment generation.
- CRScore leverages Large Language Models of Code and static analysis tools to analyze code changes and identify relevant **maintainability** issues like code smells to generate pseudo-references.
- Then it leverages sentence transformers and semantic similarity of the review sentences with the pseudo-references to compute the conciseness, comprehensiveness, & relevance of the review.
- Creation of a human-annotated code review quality dataset for comparing review quality metrics. CRScore exhibits the greatest correlation and sensitivity to human judgment (**0.95** Spearman rank correlation with human ranking of systems) compared to popular reference-based metrics like BLEU.

Skills

Coding	Python (expert), C/C++, Bash (familiar), Javascript (novice)
Frameworks	PyTorch, HuggingFace, Fairseq, NLTK, spaCy, Tensorflow, FastAPI, Flask, Django, PyQt5, Jupyterlab, OpenCV, Git

Awards

2024
- **Among the top 10 teams selected for the Amazon Trusted AI Challenge**, Awarded $250K grant to pursue research related to development of secure Large Language Models of code
- **Best paper and best student paper nominations**, Artificial Intelligence in Education (AIED 2024)

2022
- **2nd place** at Deep Learning Labs OpenAI GPT-3 Hackathon

2021
- **DAAD WISE Scholarship**
- **MITACS Globalink Scholarship**
- **Bronze** at Inter IIT Technology Meet: Member of IIT Kharagpur contingent

2019
- **2nd** place at Intelligent Ground Vehicle Competition (IGVC)

2018
- All India Rank **1248** in JEE Advanced and **1618** in JEE mains among 1M candidates
- **Kishore Vaigynaik Protsahan Yojana (KVPY) Scholarship**