Hyunjae Suh

Irvine, CA hyunjas@uci.edu hyunjas.github.io hyunjae-suh-607119146

Summary

Ph.D. student specializing in software engineering with a focus on code generated by large language models. Experienced in conducting in-depth research to enhance understanding and improve the quality of LLM-generated source code.

Education

University of California, Irvine — *Irvine, CA*
Ph.D. in Software Engineering — *Sep 2023 – Present*

Kookmin University — *Seoul, South Korea*
BS in Computer Science — *Mar 2017 – Aug 2023*

Experience

Ph.D. Student — *Irvine, CA*
University of California, Irvine — *Sep 2023 – Present*

- Conducted research in software engineering with an emphasis on LLM-generated source code.
- Investigated key characteristics of LLM-generated source code, including its accessibility and similarity to human-written code.

Graduate Research Assistant — *Remote*
Ebay — *Aug 2023 – Dec 2023*

- Worked on automated commit message generation using open-source LLMs.
- Optimized the use of LLMs to streamline automated commit message generation.

Publications

An Empirical Study on Automatically Detecting AI-Generated Source Code: How Far Are We?
Hyunjae Suh, Mahan Tafreshipour, Jiawei Li, Adithya Bhattiprolu, Iftekhar Ahmed
Accepted at the 47th IEEE/ACM International Conference on Software Engineering [ICSE 2025]

Does the Order of Fine-tuning Matter and Why?
Qihong Chen, Jiawei Li, **Hyunjae Suh**, Lianghao Jiang, Zheng Zhou, Jingze Chen, Jiri Gesi, Iftekhar Ahmed
https://arxiv.org/abs/2410.02915

Projects

Detection of LLM-generated Source Code

- Proposed techniques that include fine-tuning LLMs, leveraging embeddings, and analyzing code features to accurately identify LLM-generated code.
- Improved the detection of LLM-generated source code, achieving an F1-score of 82.55.
- Research paper accepted for publication at ICSE 2025.

Accessibility of LLM-generated Source Code

- Evaluated the accessibility of source code generated by LLMs in the web domain.
- Applied prompting techniques to generate code with LLMs, resulting in enhanced accessibility.
- Utilized accessibility evaluation tools to assess and ensure compliance with WCAG 2.1 guidelines.

The impact of Fine-tuning Order on Language Models for Software Engineering

- Developed machine learning pipelines for fine-tuning transformers with various combinations of software engineering tasks.
- Evaluated the impact of fine-tuning order on final task performance.

Teaching Experience

University of California, Irvine *Irvine, CA*
Teaching Assistant *Sep 2023 – Present*

- IN4MATX 115 - Software Testing, Analysis, and Quality Assurance
- ICS 10 - How Computers Work
- ICS 32 - Programming with Software Libraries