Moshi Wei
Linkedin: https://www.linkedin.com/in/moshi-wei/
Email: wmswms938@gmail.com Google Scholar: https://scholar.google.ca/citations?user=MoshiWei

Summary

I am Moshi, a third-year PhD student specializing in machine learning applications on source code, including code generation, API recommendation, and automated bug repair. I am a data-driven problem solver with a focus on practical implementation and continuous improvement. I consider myself a hacker, as I am target-driven and quick to learn new skills to achieve my goals.

Skills

Machine Learning, Code Generation, Code Analysis, Data Analysis, Python, Deep Learning, Natural Language Processing, Research, Problem Solving, Software Engineering, Data Visualization

Experience

Applied Scientist (Intern) – Huawei Research Canada Jan. 2023 – Apr. 2023

- Developed new algorithm based on information bottleneck theory and result in a prototype for 6G advanced hyperspeed signal transmission model, compressing the signal by 40% while maintaining the same quality using sequential data modeling techniques with RNN and LSTM. This work resulted in one patent and one published paper.
- Designed an LLM-based configuration model for equipment configuration, saving 90% of human labor costs on network equipment installation and configuration.

Research Assistant, York University Sep. 2020 – Now

- Developed new SOTA API recommendation model CLEAR for API recommendation with a 40% accuracy improvement using Contrastive learning on Sentence-Bert model.
- Replicated and fine-tuned existing SOTA algorithm for Code models such as CodeSearch net, CodeGPT, CodeBert, BIKER, and DeepAPI and extracted API usage patterns for Deep learning libraries. This work improves the performance of the SOTA by 20% with strong generalization ability.

Machine Learning Engineer – BI, Achievers May. 2019 – Sep. 2020

- Designed and developed new algorithm for the awarded collaborative filtering-based user recommendation system with Spacy, Scikit-learn, and pandas. Ran the system on 200 users for testing and feedback.
- Developed and maintained daily ETL tasks with Centerprise ETL and PostgreSQL.
- Speeded up TB-level ETL data warehouse restoration time by 24 times by analyzing and refactoring workflow with new algorithms using PostgreSQL and Python.

Project

ChatGPT Clone - Wechat payment

- Developed ChatGPT4.0 clone that accepts Wechat payment, receiving 100% positive user feedback. (LangChain, Streamlit, Python)
- https://chatgpt-wechat.streamlit.app/

TarotGPT - AI Powered Tarot reading app

- Developed algorithm for a tarot interpreter app that answers your questions based on card reading, supporting Chinese languages. receiving 100% positive user feedback. (LangChain, Streamlit, Python)
- https://tarot-wechat.streamlit.app/

ResumeLab - AI Powered One-stop Resume editor

- Developed algorithm for a resume editor app that identifies problems and polishes your resume. (LangChain, Streamlit, Python)

Publication

- **CoCoFuzzing: Testing Neural Code Models with Coverage-Guided Fuzzing**
  Moshi Wei, Yuchao Huang, Jinqiu Yang, Junjie Wang, and Song Wang
  IEEE Transactions on Reliability (Transactions on Reliability'22 Journal)

- **API Recommendation for Machine Learning Libraries: How Far Are We?**
  Moshi Wei, Yuchao Huang, Junjie Wang, Jiho Shin, Shiri harzevili Nima, and Song Wang
  ESEC/FSE 2022 (acceptance rate=22%)

- **CLEAR: Contrastive Learning for API Recommendation**
  Moshi Wei, Shiri harzevili Nima, Yuchao Huang, Junjie Wang, and Song Wang
  ICSE 2022 (acceptance rate=26%)

- **Automatic Unit Test Generation for Machine Learning Libraries: How Far Are We?**
  Song Wang, Nishtha Shrestha, Abarna Kucheri Subburaman, Junjie Wang, Moshi Wei, and Nachiappan Nagappan
  ICSE 2021 (acceptance rate=23%)

- **Yet Another Combination of IR-and Neural-based Comment Generation**
  Yuchao Huang, Moshi Wei, Song Wang, Junjie Wang, and Qing Wang
  IST 2022

- **Coconut: combining context-aware neural translation models using ensemble for program repair**
  Lutellier Thibaud, Pham Hung Viet, Pang Lawrence, Li Yitong, Wei Moshi, and Tan Lin
  ISSTA 2020

- **Survey on Automated API recommendation (in progress)**
  Moshi Wei, Song Wang.

Education

- York University, Ph.D. of Software Engineering — Sep. 2020 – Jan. 2024
- University of Waterloo, Master of Software Engineering — Dec. 2017 – May. 2019
- University of Ottawa, Computer science — Sep. 2012 – May. 2016