Zhensu Sun

Email: 87su@tongji.edu.cn
Website: https://v587su.github.io/

Research Interests

I'm a graduate student majoring in software engineering at Tongji University, China. My research interest is mining software repositories with various techniques like information retrieval and natural language processing. The goal of my research is to boost the productivity of developers.

Education Background

Tongji University, Shanghai, China *Sep 2018 - Expected Mar 2021*
M.S. in Software Engineering, School of Software Engineering
GPA: 90.9/100
Awards: 2nd Prize at National Post-Graduate Mathematical Modeling Contest (2018)

Tongji University, Shanghai, China *Sep 2014 - Jun 2018*
B.S. in Logistic Engineering, School of Transportation Engineering (with Summa Cum Laude)
GPA: 4.2/5.0 (Top 3 of 28)

Research Experiences

High-quality Training Dataset for Neural Code Search *Nov 2020 - Present*

- Proposed a filtering framework for code search datasets to bridge the gap between code comments and real queries by utilizing Variational Auto-Encoder (VAE).
- Presented a public code search dataset, containing 2.4 million comment-code pairs, which is the first systematically filtered training dataset for neural code search.

Deep Learning-based Semantic Code Search *Aug 2019 - Dec 2020*

- Proposed a deep learning-based semantic code search model, which additionally learns the contexts of code snippets.
- Demonstrated the effectiveness of code context in code search tasks using ablation experiments.

Requirement-based Library Recommendation *Dec 2018 - Aug 2019*

- Identified a new perspective for library recommendation: textual requirement description-based library recommendation for software projects, which avoids two key problems in existing models including requirement insufficiently matching and cold-start problem
- Proposed a model based on Seq2seq to recommend available third-party libraries with a SuccsessRate@10 of 90.4%, which proves the practicability of our perspective.

Publications

- **Zhensu Sun**, Yan Liu, Ziming Cheng, Chen Yang, Pengyu Che. "Req2Lib: A Semantic Neural Model for Software Library Recommendation". Published in SANER 2020 (CCF-B).
- **Zhensu Sun**, Yan Liu, Chen Yang, Yu Qian. "PSCS: A Path-based Neural Model for Semantic Code Search". Arxiv.

Skills

- **Programming Languages**: Python, JavaScript, SQL
- **Frameworks**: Pytorch, Pandas, Numpy, Scikit-Learn, Node.js, React.js
- **English**: IELTS: Listening: 7.5, Reading: 8.5, Writing: 6.5, Speaking: 6.0, Overall Band Score: 7.0