Yuxiang Zhu

zzzzzyx.github.io | (+86)18938889050 | zyx@smail.nju.edu.cn

Education

Software Institute, Nanjing University — Nanjing, China
Master of Engineering in Software Engineering; GPA: 88.4/100 — *Expected Jun. 2020*

Software Institute, Nanjing University — Nanjing, China
Bachelor of Engineering in Software Engineering; GPA: 85.6/100 — *Sept. 2014 - Jun. 2018*

Academic Experience

Research Assistant — Nanjing, China
Software Engineering Group, State Key Laboratory for Novel Software Technology, NJU — *Nov. 2016 - Present*

- **Issue Classification**: Developed a model to **classify issues** from GitHub and Jira.
  - Collected and preprocessed over 1.2M issues in GitHub and Jira.
  - Altered and implemented the k-NN algorithm to **detect issues with misclassified input**.
  - Implemented and trained an attention-based bi-LSTM, resulting in **11.1% improvement in F1-score** compared to SVM model, and **5%-8% improvement in F1-score** compared to state-of-the-art approaches.
  - **First-authored** a paper on the findings: **Y. Zhu** *et al.*, "A Bug or a Suggestion? An Automatic Way to Label Issues," arXiv:1909.00934 [cs], Sept. 2019.
- **Pull Request Summarization**: Implemented and trained a **seq-to-seq model** to summarize diff files in pull requests into short informative messages. Achieved a BLEU score of 17.3 on a test set of 30k pull requests.
- **Program Comprehension**: Conducted a **systematic literature review** in the field of program summarization.
  - Examined 41 relevant studies from 2010 to 2019. Compared different approaches and discussed different techniques used in the state-of-the-art approaches.
  - **First-authored** a paper on the findings: **Y. Zhu** and M. Pan, "Automatic Code Summarization: A Systematic Literature Review," arXiv:1909.04352 [cs], Sept. 2019.
- **Requirement Extraction**: Developed a tool to automatically collect and **analyze user reviews** in Google Play.
  - Collected, filtered, and preprocessed 394k user reviews from 100 popular apps.
  - Designed and implemented the K-Means++ algorithm to **cluster reviews by tf-idf word vector**, and then selected words with highest tf-idf weight in each cluster as **keywords**.
  - Created a completely automatic tool and developed a Java Web website to display the result.
- **Stars' Microblog Analyzer**: Designed a cluster algorithm to **group stars who posted similar content** on Weibo (known as Chinese Twitter). Used the **PageRank** Algorithm to identify the central star of each group.

Professional Experience

Software Engineer Intern — Shanghai, China
Meituan-Dianping Inc. — *Jul. 2017 - Jan. 2018*

- **GroupJoy**: As the project manager, **led a team of six newcomers** to develop a platform for internal employees to launch and attend activities, socialize, and find common interests. We spent four weeks on **collecting requirements, designing and coding**; currently, our system is **still serving** thousands of employees.
- **CustomerInfoService**: **Maintained and iterated** a large Java Web System full-stack-ly, which enables more than 40k sales staff to register, manage, and contact merchants all over China.

Scholarships

- First-Class Scholarships for Outstanding Students, Nanjing University (Top 10%) — *Nov. 2016*
- Second-Class Scholarships for Outstanding Students, Nanjing University (Top 20%) — *Nov. 2015*

Skills

- **Languages**: Python, Java, SQL, C++, PHP, HTML, CSS
- **Technologies**: Maven, Git, Linux, Microsoft Office, Spark, LaTeX
- **Development**: Object Oriented, Databases, Machine Learning, NLP
- **Standard Tests**: TOEFL 107; GRE V160+Q170+AW3.5