

Yuxiang Zhu

[zzzzzyx.github.io](https://github.com/zzzzzyx) | (+86)18938889050 | zyx@smail.nju.edu.cn

EDUCATION

Software Institute, Nanjing University
Master in Software Engineering; GPA: 88.4/100

Nanjing, China
Expected Jun. 2020

Software Institute, Nanjing University
Bachelor in Software Engineering; GPA: 85.6/100

Nanjing, China
Sept. 2014 - Jun. 2018

ACADEMIC EXPERIENCE

Research Assistant

Software Engineering Group, State Key Laboratory for Novel Software Technology, Nanjing University

Nanjing, China
Nov. 2016 - Present

- **Issue Classification:** Developed a model to classify issues from Issue Tracking Systems such as GitHub and Jira
 - * Collected and preprocessed over 1.2M issues in GitHub and Jira
 - * Altered and implemented the k-NN algorithm to detect issue that is wrongly classified by its submitter
 - * 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
 - * **Y. Zhu et al.**, "A Bug or a Suggestion? An Automatic Way to Label Issues," [arXiv:1909.00934](https://arxiv.org/abs/1909.00934), Sept. 2019.
- **Pull Request Summarization:** Implemented and trained a sequence-to-sequence model to summarize diff files in pull requests to short informative messages. Achieved BLEU score of 17.3 on test set of 30k pull requests
- **Program Comprehension:** Conducted a systematic literature review in the field of automatic source code summarization
 - * Examined 41 relevant studies from 2010 to 2019. Compared different approaches and discussed different techniques used in the state-of-the-art approaches.
 - * **Y. Zhu** and M. Pan, "Automatic Code Summarization: A Systematic Literature Review," [arXiv:1909.04352](https://arxiv.org/abs/1909.04352), 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. Then selected words with highest tf-idf weight in each cluster as keywords.
 - * Formed a completely automatic tool and developed a Java Web website to display the result

PROFESSIONAL EXPERIENCE

Software Engineer Intern

Meituan-dianping Inc.

Shanghai, China
Jul. 2017 - Jan. 2018

- **GroupJoy:** As the project manager, led a team of six newcomers to design and develop a platform for internal employees to launch and attend activities, to socialize and find common interest. We spent 4 weeks on collecting requirements, designing software architecture and coding. Now, our system is still serving thousands of staffs
- **CustomerInfoService:** Maintained and iterated a Java Web System full-stack-ly, which helps more than 40k sales staff register, manage, and contact with merchants all over China.
- **Data Analysis:** Analyzed and visualized data of merchant relationship, involving over 1M merchants. Independently set up a small system which emails every day statistics graphics about our merchant data to all(15) colleagues in our group

PROJECTS

- **Stars' Microblog Analyzer:** A model to analyze Chinese entertainment stars' behavior by microblog content
 - Collected 60k microblogs of 500 stars in Weibo. Designed and implemented a cluster algorithm to group stars that send similar content. Used the Pagerank Algorithm to identify the center star of each group
- **Stock Analysis:** A website that comprehensively and visually displays statistics and graphics of stock data
 - Crawled historical records of stocks from famous stock website and generated statistics and graphics
 - Built a sub-system that enables user to design static buy-and-sell strategy and simulate it to see its profitability

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, Javascript
- **Technologies:** Maven, Git, Linux, Microsoft Office, Spark, LaTeX
- **Development:** Object Oriented, Relational databases, Machine Learning, NLP
- **Standard Tests:** TOEFL 107; GRE V160+Q170+AW3.5