Junming Cao

junmingcao@fox[illegible]n (+86)15221928719

Education

M.Eng in Software Engineering | Shanghai Jiao Tong University 2018.9 -
Big data software engineering lab | Advised by **Beijun Shen** (http://base.sjtu.edu.cn/~bjshen/) Shanghai, China
Research direction: Bug Localization, Code Knowledge Graph

B.Eng in Software Engineering | Shanghai Jiao Tong University 2014.9 - 2018.6
Main courses: Data Structures and Algorithms / Software Engineering / Data Mining / Computer System Shanghai, China

Research

BugPecker: Locating faulty methods with deep learning on revision graphs

Junming Cao, Shouliang Yang, Wenhui Jiang, Hushuang Zeng, Beijun Shen, and Hao Zhong.

The 35th IEEE/ACM International Conference on Automated Software Engineering (ASE 2020, new idea)

CocoQa: Question Answering for Coding Conventions over Knowledge Graphs

Tainjiao Du, **Junming Cao,** Qinyue Wu, Wei Li, Beijun Shen*, Yuting Chen.

The 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019, tool demo)

Constructing a Knowledge Base of Coding Conventions from Online Resources

Junming Cao, Tianjiao Du, Beijun Shen*, Wei Li, Qinyue Wu, Yuting Chen.

The 31th International Conference on Software Engineering and Knowledge Engineering(SEKE 2019)

Project

BugPecker | Bug localization system 2019.9 - 2020.7
Big Data Software Engineering Lab, Shanghai Jiao Tong Universit Shanghai, China

- **Challenge:** Lots of methods are short and lack sufficient details to be matched against bug reports.
- **Approach:** Build revision graphs from past fixes, and discover relations inside them to expand the details ~~for~~ methods ~~and calculate various features~~ to assist the localization.
- **Result:** ~~Compared with two recent approaches,~~ BugPecker improves the MAP values of ~~all three projects by five times.~~

CocoQa | Question answering system for coding conventions 2018.9 - 2019.5
Big Data Software Engineering Lab, Shanghai Jiao Tong University Shanghai, China

- Parse the question into a SPARQL query and acquire candidate nodes and relations from knowledge graph by subgraph matching.
- Consider coding conventions as paragraphs, then retrieve answers with Match-LSTM model.
- Apply a logistic regression classifier to rank the candidates.
- Implemented CocoQa as an Intellij Idea's plugin and ~~as well~~ a web service.

DeepEye | Pornographic image classification system 2017.7 - 2017.9
Backend Developer Intern, Youtu AI Lab, Tencent Shanghai, China

- Develop~~ed~~ efficient backend services by deploying new modules of Nginx.
- Develop~~ed~~ a distributed crawling system with Scrapy. Crawled millions of images per day ~~for~~ porn image identifying algorithm ~~training and testing~~.

Skills

Java / C++ (intermediate)
Go/Python/Scala(begin)
Git / Spark / PyTorch

Honors

- CUMCM (China Undergraduate Mathematical Contest in Modeling 2018): National Second Prize in 2018
- MCM(Mathematical Contest in Modeling): Meritotious Winner in 2017