Xiang YUE

Dorm B638 Bldg. 3, No.163, Xianlin Avenue, Nanjing City, China 210046

+86 157 2061 8301 | xiangyue@smail.nju.edu.cn

EDUCATION

Nanjing University (NJU) **Nanjing, China**

B.Eng. in *Experimental Class of Fin-Tech* ***Sep.2016 – Jun.2020 (Expected)***

- GPA: **4.53/5.0 Rank: 3/20**
- **Experimental Class of Fin-Tech** is a **Pilot Class** in Nanjing University to focus on the combination knowledge learning in the areas of **Computer Science**, **Mathematics** and **Finance**.
- **Core courses:** *Introduction to Computer Systems, Foundation of Computer Networks, Big Data Processing Technology, Operating System, Financial Economics, Corporate Finance*

RESEARCH EXPERIENCE

Interest: **Natural Language Processing**, **Machine Learning**, **Data Mining**

Natural Language Processing Group, Nanjing University

Undergraduate Research Assistant, supervised by Prof. Shujian Huang ***Oct.2018 – present***

- **Topic:** *Combining Character and Word Information in Neural Machine Translation based on a Multi-Level Attention.*
- Proposed an encoder with character attention to augment the (sub)word-level representation with character-level information, given the fact that neural machine translation models face with difficulties to embed sentences in a certain granularity.
- Managed to devise a decoder with multiple attentions that covered both sentences and characters to control translation to process better.
- Created the model with incorporating "Char-att" encoder as well as "Multi-att" decoder, successfully outperforming the models of baseline in experiments.

Automatic Identification of Outdated Requirements by Using Closeness Analysis based on Source Code Changes, Nanjing University

Innovation Research Project, supervised by Dr. Hongyu Kuang ***Sep.2017 – Sep.2018***

- Utilize the "closeness" for each call dependency between two methods, to refine the quality of the keywords extracted from the source code changes, and thus improve the accuracy of outdated requirement identification.
- Succeeded to capture the relevant code changes with accuracy higher than 86%, sensitivity (TPR) higher than 80%, and specificity (TNR) higher than 68%.

Generation of Short Summaries of Commits by Using Closeness Analysis based on Source Code Changes and Nearest Neighbor Generation Techniques, Nanjing University

Undergraduate Research Assistant, supervised by Dr. Hongyu Kuang ***Oct.2018 – Present***

- According to the large application of NLP methods, we consider generating short summaries of commits of Git projects by combining our Closeness Analysis Technique with the Nearest-Neighbor Generation (NNGen) Technique – the machine learning technique, to save the maintainers' time to write more precise commit messages when committing.

PROJECT EXPERIENCE

Generator for Structural Change Groups in different project's versions using Closeness Analysis based on Source Code Changes

Designer, Developer ***Sep. 2017– Nov.2018***

- An automatic tool developed by Java, which is useful for generating change groups in different project's version, using the closeness analysis technique.
- Can be applied in any realistic projects that published in GitHub. Efficient to show the change groups that represent the structural differences of certain two different commits or version of project.

Crazy Whale – Independent Parkour Game

Designer, Developer ***Apr. 2017– Jul.2017***

- An independent parkour game written by C++ with graphical library. Players play the game in the role of an NJU whale to run and avoid obstacles to survive. The game simulated and generated the gravity effect, collision detection and random algorithm for obstacle generation.

Non-dictionary-based Chinese Input Method

Designer, Developer ***Oct. 2017– Dec.2017***

- A Chinese input method developed by C++. In Chinese, an input method needs two things to confirm an input: a dictionary base and a correct split of a pinyin string. But it can be tricky when there is no dictionary.
- Solved the problem by grabbing the N-gram-string from 4GB corpus in internet, and generated possible words using Naive Bayes. Made it to cut the pinyin string and display the final input choices.

HONOR & AWARDS

- **People's Scholarship** – First Prize, Nanjing University (**3/35**) ***Jun.2017***
- Programming Design Contest – Second Prize, Nanjing University ***Oct.2017***
- Mathematical Contest in Modeling – Honorable Prize, COMAP ***Apr.2018***
- Citi Financial Innovation Application Competition – Third Prize, Citi Bank ***Nov.2018***
- 2018 SAS Data Mining Champion – First Prize, SAS China (Top 30) ***Dec.2018***
- Excellence Scholarship – Second Prize, Nanjing University ***Jan.2019***

SKILLS

Programming: C99, C++11, Python, R, C#, Java, Scala, Intel x86 Assembly, MATLAB, Shell
Platform: Windows, Linux