Pengyu Li

E-mail: lipy12@mails.tsinghua.edu.cn

Tel: 00-86-13051308960

Department of Physics in Tsinghua Univ., 100084, Beijing

Personal Git Repositories: github.com/RyuLee

EDUCATION

Fundamental Mathematics and Physics, Department of Physics, Tsinghua University

Bachelor of Science   Anticipated Graduation: July 2016

Anticipated Courses:

Data Structure and Algorithm (85), Introduction to Machine Learning(93), Big Data System(99), C++ Programming in Linux etc.

Professional Skills: Java, Scala, C++, Python, R, MATLAB, SQL, Machine Learning, Data Mining.

SELECTED WORK

Student Online Management System

Programmer   Aug, 2014

- Used Java to build serve page and JavaScript to implement many functions. Designed the logic between functions and database. Built a whole system involving BBS, curriculum enrollment and information management on Web.

Information Retrieve Program

Designer and programmer   Nov-Dec, 2014

- Captured the web pages and extracted Chinese contents on pages. Designed the algorithm to use context of the content extracted and Chinese word dictionary to divide sentences to separate words.
- Used Jaccard Similarity to evaluate the object page's most similar pages in database and return the results.

Hospital Medicine Management System

Designer and programmer   Dec, 2014

- Designed the system logic and implement it using Oracle database. Finally accomplished a software with pretty GUI.

State Key Laboratory of Intelligent Technology and Systems

Research Assistant   Dec, 2014 - Jun, 2015

- Work in Prof. Min Zhang's group and study on how to improve performance of state-of-art recommendation algorithms by using *Learning to Rank*'s methods.

WING Group in National University of Singapore, School of Computing

Research Assistant   Jul, 2015 - Sep, 2015

- Work in Web Information Retrieval/ Natural Language Processing Group under Prof. Kan Min Yen's supervision. Develop the Wikifier System aim to help MOOC users to get better experience when they use discussion forum in MOOC platform.
- Use NLP technique and regular expression to find terms in discussion forum which have entities in MOOC System like lecture element, slides elements. Then link them with entities.
- The difficulties in the wikify process is to disambiguate and figure out which term has real entities on the system. I use NLP algorithm and build corpus with extracted data.

ADDITIONAL

Language: Proficient in **Mandarin**, **English**, **Japanese** as hobby.

Volunteering and other student clubs: As a leader in advertise group in department student union for one year. As a volunteer, joined the *Red Cross* in school and taught poor children computer skills in *Xiaojiahe primary school* which locates in remote area in Beijing for a semester.