Yifu Wang

PHONE: (626) 226 6282 EMAIL: yifuw@cmu.edu
GITHUB: https://github.com/wangyifu LINKEDIN: https://www.linkedin.com/in/yifuwang

EDUCATION

08/14 - 05/16 **M.S. Computer Science**, Carnegie Mellon University, School of Computer Science
Courses: Advanced Data Structure and Algorithms, Machine Learning,
Cloud Computing, Big Data System in Practice, Search Engine
Teaching: Data Structures for Application Programmers, JAVA for Application Programmers

09/10 - 06/14 **B.S. Software Engineering**, Dalian University of Technology
Courses: Operating System, Compiler Techniques, Database Systems,
Computer Network, Computer Architecture, Data Structure and Algorithms

WORK EXPERIENCE

05/15 - 08/15 Software Engineer Intern at *Quixey, Inc.*

- Implemented Chinese query tokenizer based on **Early parser**.
- Using **chart parse** to store intermediate result.
- Applied **dynamic programming** in candidate path elimination.
- **40% faster** and **1.6% gain of DCG score** compared with current tokenizer.
- Integrated into Quixey app search and auto suggestion in China market.
- Intellectual property application has been accepted and in process.

08/13 - 02/14 SDET Intern at *VMware, Inc.*

- Implemented **image comparison plugin** of the automation testing framework using **C#**
- Developed several web applications based on **LAMP(Linux, Apache, MySQL, PHP)** stack.
- Visualize product development progress, employee KPI, and bug report.

PROJECTS

BIO-ASK:
- Question answering system in medical field based on **UIMA architecture**
- Implemented document retrieval algorithm based on **Stanford NLP** and **Lingpipe**
- Developed result evaluation system based using **BM25** and **GMAP**

DOTA2 COUNTER-PICK:
- Hero recommendation engine for Dota 2 game in python.
- Generated feature matrix from 30000 matches using **urllib, numpy, pandas**.
- Suggest heroes based on machine learning algorithms including **logistic regression**, **K-nearest neighbors**, **SVM with RBF kernel** using **sci-kit** package.
- Achieved **70% accuracy** in predicting match outcomes.

TARTAN ENGINE:
- Simple search engine based on **lucene**
- Supports both **unstructured** and **structured** query.
- Implemented operators: **AND, OR, NEAR, WINDOWS, SYN**.
- Implemented several retrieval models: **BM25, Indri, Ranked Boolean**

DRAW SOMETHING:
- Online graphic multiplayer game runs on **Linux**
- Developed **graphic** part and **network communication** part using **C++** and **Qt**
- Used **bezier curve** to represent and serialize canvas.
- Implemented **multi-thread TCP server** using Qt to support concurrent connections.

SKILLS

Languages: JAVA, Python, C++, MySQL, shell script, Javascript, PHP, Scala, Matlab, R

Other Technologies: Web Development(LAMP), Linux, AWS, Guava, Hadoop, ElasticSearch, Mongo, Mahout, Git, Maven, Gradle, Docker