

Yuchen Cao

yuchenc1@cs.cmu.edu · 412-478-9381 · <https://www.linkedin.com/in/yuchen-cao-cmu> · <https://github.com/yuchenc1>

EDUCATION

Carnegie Mellon University, School of Computer Science, Language Technologies Institute
Master of Science in Biotechnology Innovation and Computation

Pittsburgh, PA
Sep 2017 - May 2019

Zhejiang University, Mathematics Department
Bachelor of Science in Applied Mathematics

Zhejiang, China
Sep 2013 - Jul 2017

PROFESSIONAL EXPERIENCE

Snowflake Computing | Service Runtime Team
Software Engineer

San Mateo, CA
Aug 2018 - Present

Uber Technology | Software Network Team
Software Engineer Intern

San Francisco, CA
May 2018 - Aug 2018

- Built the staging environment for discovery, health check and traffic routing services using Go. Prototyped a custom dns resolver watching updates to staging zookeeper.
- Improved the lint check tool for traffic group configuration of all services at Uber using Python, making the jenkins check and related service 10 times faster.
- Implemented resolver and balancer using grpc for Ubers own naming service client.

Yitu Technology | Developing Engineer Team
Software Engineer Intern

Shanghai, China
Feb 2017 - Jun 2017

- Developed GPU computing in face retrieval and comparison using C++.
- Implemented LevelDB in car recognition platform, which supports approximate matching with more than 300 QPS.
- Applied FFmpeg into transferring video stream for real time camera monitoring.

PROJECTS

Risk-oriented Relational Database

Aug 2018 - Dec 2018

- Implemented extendable hash-table, lru cache, buffer pool manager, b plus tree index, lock manager with two phase locking, log manager and recovery for a relational database.

Twitter Analytics Web Service

Feb 2018 - May 2018

- Built a high-performance multi-tier web service on AWS with limited budget using JAVA handling more than 20k RPS.
- Implemented ETL on JSON Twitter data set (1 TB) using MapReduce on GCP.
- Designed schema and optimized MySQL and HBase databases to support large scale and strong consistent read/write query under multithreading design.

Lung Nodule Analysis

Feb 2018 - May 2018

- Developed an automatic lung nodule detection algorithm using 888 CT scans from LUNA 2016 dataset.
- Implemented the actor-critic algorithm for reinforcement learning framework for non-greedy decoding dependency parsing.
- Loaded and reprocessed unbalanced raw data using SimpleITK.
- Modified C3D neural network using 8 3D convolutional and 6 pooling layers using PyTorch.
- Trained the model in 30 epochs and delivered 92% accuracy on the test data.

Malloc, Shel, Proxy Labl

May 2017 - Aug 2017

- Implemented my own malloc function with high utility and throughput.
- Realized a tiny shell supporting job control and I/O redirection.
- Developed a small http proxy with cache and multithred handling SIGPIPE.

Concussion Detection Using Temporal Analysis of Speech

Jul 2016 - Aug 2016 *Research Exprience with Prof Christian*

Poellabauer in University of Notre Dame

- Took speech recordings from athletes and extracted temporal metrics from speech data using Python, CMUSphinx, Kaldi.
- Applied machine learning analysis with Logistic Regression and Neural Network into finding patterns that are indicative of a concussion using Matlab with the accuracy of 96.05%.

SKILLS

Languages: GO, JAVA, Python, C++/C, Bash, Matlab, SQL, R, LaTeX

Frameworks: Linux, Git, Zookeeper, Kafka, GRPC, MapReduce