Yangcheng Gu

Carnegie Mellon University | yangcheg@andrew.cmu.edu | (412) 954-8944

https://www.linkedin.com/in/yangcheng-gu/ | https://ferv3455.github.io/

Education

Carnegie Mellon University, Master of Science in Information Networking GPA 4.00/4.00 | August 2024 – Present

- *Courses:* Introduction to Computer Systems, Networking and the Internet, Distributed Systems, Embedded Systems
- *Expected graduation:* December 2025

Tsinghua University, Bachelor's in Software Engineering GPA: 3.90/4.00 | August 2019 – July 2024

- *Courses:* Computer Architecture, Database Systems, Data Structures and Algorithms, Artificial Intelligence
- *Honors:* Scholarship for Academic Excellence (four times in 2020, 2021, 2022, 2023)

Work Experience

Algorithm Application Engineer (Intern) June 2023 – August 2023

Risk Control R&D Department, Beijing Jingdong Century Trading Co., Ltd.

- Operated on Apache Hadoop and Spark to compile a dataset and fitted a pre-trained click-farming detection model to it
- Devised a specialized image similarity algorithm based on partial analysis with statistic methods and feature matching, yielding an average accuracy of over 95% on synthetic datasets and over 98% on real ones
- Developed, optimized and deployed a click farming detection pipeline in online production environment with the proposed image similarity algorithm, able to process up to 300,000 images per day and spot almost 100% of tested real-life anomalous comment images

Publications

Deep Active Learning with Noise Stability (AAAI-24 conference) February 2024

Xingjian Li, Pengkun Yang, ***Yangcheng Gu***, Xueying Zhan, Tianyang Wang, Min Xu, Chengzhong Xu

Research Projects

Bi-level Optimization for Inductive Transfer Learning July 2023 – July 2024

Computational Biology Department, Carnegie Mellon University

- Proposed and implemented a model pre-training method where sample weights are assigned with a neural network learned with DARTS, in order to boost its performance in transferring to a differently distributed dataset
- Assessed its performance with multi-domain samples and achieved a 3% increase in accuracy with a limited dataset
- Merged the framework with SimCLR self-supervised learning architecture and investigated sample re-weighting in domains of self-supervised deep learning with both natural and medical datasets

Log Data Encoding for Efficient Storage in Apache IoTDB October 2022 – May 2023

School of Software, Tsinghua University

- Introduced advanced single-character operations and generalized string edit distance as a string encoding scheme
- Performed character re-weighting based on their frequencies of occurrence and delivered higher space efficiency
- Classified log data according to string formats with clustering algorithms, minimizing required data storage
- Optimized the algorithm temporally with a linear-time cosine distance algorithm for strings based on q-grams

Course Projects

The Raft Consensus Algorithm, Distributed Systems October 2024 – November 2024

- Implemented the Raft consensus algorithm with Go, involving leader election and maintaining log consistency
- Skills and tools: Go, concurrent programming, distributed algorithms, consensus protocols

MarsOJ Online Coding Competition Platform, Software Engineering September 2022 – January 2023

- Created and maintained modules, structures and interfaces of the frontend framework as a group co-leader
- Completed real-time web communication functions with Socket.IO in frontend and backend projects
- Skills and tools: Javascript (Vue.js), Python (Flask), HTML/CSS, WebSocket (Socket.IO)

Simple C++-LLVM Compiler, Principles of Assembly and Compilation November 2022 – January 2023

- Designed a compiler from scratch with Python translating C++ code to LLVM IR, which can be executed with LLVM
- Skills and tools: Python, C++, LLVM IR, compiler principles, finite automata

Summary

- **Proficient in:** Python, C/C++, Go, Java, JavaScript, HTML/CSS, SQL (MySQL)
- **Familiar with:** Hadoop, Spark, Docker, PyTorch, OpenCV, Qt, Vue.js, Linux, Android, LaTeX, Markdown
- Currently looking for Summer 2025 internship opportunities (able to work **from May 12 to August 22**)