GU Zhenhao

 (+65) 8551-4957 | ✉ guzh@nus.edu.sg | ⓖ github.com/GZHoffie | ⓛ linkedin.com/in/zhenhao-gu

Education

National University of Singapore   *Singapore*

Master of Computing, Computer Science Specialisation   *Aug 2021 - Current*

- CAP: 5.0/5.0
- **Relevant Courses:** *Algorithms at Scale* (A), *Uncertainty Modelling in AI* (A+), *Distributed Systems* (A), *Combinatorial Methods in Bioinformatics* (A), *Advanced Algorithms*, *Advanced Topics in Computational Biology* (A+)

Shanghai Jiao Tong University   *Shanghai, China*

Bachelor of Science in Electrical and Computer Engineering, Minor in Data Science   *Sep 2017 - Aug 2021*

- GPA: 3.62/4.0   Rankings: 15/164
- **Relevant Courses:** *Probabilistic Methods in Engineering* (A+), *Data Structures and Algorithms* (A), *Intro to Artificial Intelligence* (A+), *Intro to Algorithms* (A+), *Linear Algebra* (A+)

Work Experience

Department of Computer Science, National University of Singapore   *Singapore*

Teaching Assistant   *Aug 2022 - Current*

- Conducted tutorials and designed exam questions for CS2040S, *Data Structures and Algorithms*, by Dr. Ket Fah Chong.

UM-SJTU Joint Institute, Shanghai Jiao Tong University   *Shanghai, China*

Teaching Assistant   *Sep 2018 - Aug 2021*

- Worked as a teaching assistant for 11 courses including high-level courses such as *Intro to Operating Systems*
- Proposed innovative new ways of teaching recitation classes, e.g. interactive, self-practice quizzes on CANVAS
- Developed a popular archive of interactive *Methematica* notebooks for visualization of concepts in probabilistic methods

Intel Corporation   *Shanghai, China*

Deep Learning Software Intern   *Jan 2021 - Jun 2021*

- Improved analytics-zoo, an open-source platform using Apache Spark that deals with Machine Learning with Big Data
- Identified and removed bugs, developed new features in the source code of TensorFlow zoo orca estimator, a key API
- Explored ways of optimizing deep learning with Intel CPUs using jemalloc, OpenMP libraries and distributed data parallel plugins, and developed an end-to-end tool orca-lite-poc that accelerates PyTorch with these optimizations (at most a 700% speed up with 4 processes than pure PyTorch)

Research Experience

BucketMap: Fast Cache-efficient Bit-parallel Alignment-free Read Mapper   *Singapore*

Instructor: Dr. Wong Limsoon, School of Computing, National University of Singapore   *Aug 2021 - Current*

- Proposed a new hierarchical mapping method for finding the locations of short reads on a long DNA reference genome
- Designed a new data structure *fault-tolerant filter* that allows fast bit-parallel pop count for multiple bit arrays, facilitating fast candidate region filtering
- Reduced the sizes of index files so that they can fit in the cache of ordinary computers, and significantly speed up mapping
- Implemented a short read mapper that can find candidate regions for 1 million short reads (150 bp) in just 40 seconds
- **Stage of code**: GitHub/bucket-map

Sublinear-space Sketching Algorithm for Identifying Correlation in Stream of Samples *Singapore*
Instructor: Dr. Seth Gilbert, School of Computing, National University of Singapore *Aug 2021 - Dec 2021*
- Designed a novel, sublinear-space algorithm to estimate the ℓ_2-independence between two random variables based on a stream of samples from their joint distribution
- Mathematically proved and practically verified the bias and variance of our new estimator
- Implemented the algorithm and showed that the novel algorithm has similar or better performance than state-of-the-art sketching algorithms while practically using 16x less space
- **Publication**: arxiv/2211.10137,   **Stage of code**: GitHub/independence-in-streams

Greedy Approximate String Matching Algorithm *Shanghai, China*
Instructor: Dr. Hongyi Xin, UM-SJTU Joint Institute, Shanghai Jiao Tong University *Mar 2021 - Current*
- Designed a greedy, bit-parallel algorithm for the approximate string matching problem to eliminate the backtracking step and alignment score calculation in the traditional dynamic programming algorithms
- Proved the theoretical optimal performance (span and work) of distributed versions of existing string alignment algorithms
- Utilized Intel SIMD instructions to achieve about 5 16x speed up than the state-of-the-art distributed Needleman-Wunsch implementation
- **Stage of code**: GitHub/approximate-string-matching

Research Assistant: Analysis of Car Repair Data *Shanghai, China*
Instructor: Dr. Youyi Bi, UM-SJTU Joint Institute, Shanghai Jiao Tong University *May 2020 - Sep 2020*
- Performed data cleaning, standardizing, analysis, and time series predictions on a car repair dataset from 4S stores
- Used Chinese Natural Language Processing and TextRank algorithm to extract the most important keywords from the fault description of cars and categorize the data based on the keywords
- Analyzed the distribution of data, utilized correlation analysis, linear and logistic regression, SVM, and neural networks to predict the type of service and working hours based on the car information (brands, age, etc.)
- Adjusted hyper-parameters of the neural network so that predictions on working hours are accurate (with error less than 12 hours) 94% of the time

Interactive Search Based on Deep Reinforcement Learning *Shanghai, China*
Instructor: Dr. Paul Weng, UM-SJTU Joint Institute, Shanghai Jiao Tong University *Apr 2019 - Sep 2020*
- Predicted user activity (clicking on a hotel, interacting with a picture, or adding a filter) on the Trivago hotel recommendation website on weekends based on the user history on weekdays
- Constructed a deep reinforcement learning approach of recommendation and built a user model and simulator based on the data
- Applied the Trust Region Policy Optimization and Deep Q Network algorithm to achieve a click-through rate (probability of successful recommendation) of about 0.37
- **Publication**: arxiv/2012.06052

Skills

Programming	C, C++, Python, Shell (Bash/Zsh), R, Julia, Java, MATLAB, C#
Tools	MapReduce, MarkDown, Spark, Drill, SQLite, TensorFlow, Git, Mathematica, LaTeX, Git, Unity
Soft Skills	Teaching, Project Management, Teamwork, Problem-solving, Documentation

Awards

2021	**Honors Research Program**,	UM-SJTU Joint Institute
2021	**Annual Excellent Teaching Assistant Award**,	UM-SJTU Joint Institute
2020	**Undergraduate Excellent Scholarship**,	Shanghai Jiao Tong University
2020	**Third Prize**,	First Foxconn Industrial AI Big Data Challenge From Theory to Practice
2019	**Undergraduate Excellent Scholarship**,	Shanghai Jiao Tong University
2019	**Third Prize**,	Contemporary Undergraduate Mathematical Contest In Modeling

References available upon request.