

Yukun Zheng

+1 614-599-5385 | yzheng437@gatech.edu | GitHub: YukunZheng 2445

EDUCATION

Georgia Institute of Technology, College of Computing

Aug 2020 - May 2022

B.S. in Computer Science, Thread: Theory-Info Internetworks

Atlanta, GA

- GPA: 4.0/4.0
- **Dean's List** Spring 2021, **Faculty Honors** Fall 2020
- **Related Coursework:** Math 4032-Combinatorial Analysis, CSE 6643-Numerical Linear Algebra, CSE 6220-High Performance Computing, CS 4540-Advanced Algorithms

Ohio State University

Aug 2018 - May 2020

Majored in Math (Honor) & CSE (Honor) before transferring to Georgia Tech

Columbus, OH

- GPA: 3.989/4.0
- **Dean's List** all terms from Autumn 2018 to Spring 2020

RESEARCH EXPERIENCE

Write A Letter - A Mobile-compatible, Web-based Application for Pre-K Letter Handwriting

Aug 2021 – Present

Developer (with team of 5), Junior Design Project with client: hearatale.org

Atlanta, GA

Objective:

- Conducted user research and competitive product analysis to extract interactive features to design graphic user interface; set up sprints and iterations and customer communication
- Implemented log-in & authentication system based on 3 user groups (kids, parents, and teachers), setting up features incl. interaction between user groups, data synchronization, scoring and data analysis, etc.
- implemented features incl. trackpad/mouse/hand-on-tablet hand writing detection using OCR technology

Georgia Tech Database Group, Georgia Institute of Technology

Aug 2020 - Jan 2021

Undergraduate Research Assistant, Advisor: Prof. Joy Arulraj

Atlanta, GA

Objective: perform fuzzing using TPC-C benchmark on latest versions of 4 sql databases: MySQL, SQLite3,

CockroachDB, PostgreSQL

- Installed test versions of databases in Linux Ubuntu 18.04 simulated by vagrant virtual machine
- Independently wrote Python script to automatically run 10k+ suspicious queries on 4 databases; compared outputs and runtimes, documenting irregularities (e.g. 10+ times the difference in runtime)
- Used query minimization methods to narrow down scope of performance bugs; produced bug reports for DBMS community for future research and development

Undergraduate Research, Georgia Institute of Technology

Jan 2021 - May 2021

Research Assistant, Advisor: Prof. Qirun Zhang

Atlanta, GA

Objective: Optimization of YARPGEN (Yet Another Random Program Generator) for C/C++ compiler testing

- Gained thorough understanding of YARPGEN's principles and implementation, installed and debugged in the vagrant virtual machine
- Modified YARPGEN code to reduce runtime and focus on testing of GCC and Clang; ran 20k+ programs with YARPGEN and documented 10 suspicious programs for analysis in report to advisor

PROFESSIONAL EXPERIENCE

Internship, American Solution Inc

Dec 2019 - Feb 2020

Data Analyst Intern

Rancho Cucamonga, CA

- Wrote scripts in Java and Python to automate process of digitalization of 5 years of paper-based clients' income tax return documents; independently built a SQL database (variables: income, expense, types of deductive, etc.)
- Designed, trained machine learning model to predict tax returns based on existing data on account
- Implemented prototype of a client-facing tax planning system for data query, visualization, and analysis
 - ✱ Created user authentication and management system; embedded R and PowerBI features for user-responsive visualization; integrated prediction model with parameter-tunable interface

Zhong Yin Law Firm, Bankruptcy and Reorganization Department

Legal Intern, Industry & Market Research Intern

July 2021 - Aug 2021

Beijing, CN

- Wrote due diligence report on cloud computing company **Easted** and **Macrolink** per request of clients; investigated company's legal and financial background via internal government online database
- Conducted industry, political, and economic research and analysis on target company to advise senior staff members, evaluating potential and competitiveness of company's main tech products
- Identified, researched, and assessed various issues on-call and collaboratively worked with clients

ACADEMIC PROJECTS

Course Projects for CS4440 - Emerging Database Technologies

[1] Digital Note Management System with Dynamic Autocomplete (with team of 4)

- Implemented features incl. user authentication, note modification, search (generic & ordered), and sorting
- Trained autocomplete algorithm with pre-populated autocomplete database with n-gram web-scraping notes as training data; created user-based feature to dynamically add in words to unique page data pools
- Comparatively researched NoSQL databases to find most compatible option (for intended form of storage & authentication); applied Firebase Realtime database for real-time server-side database synchronization
- Built key-and-cursor-based front-end interface with Vue.js (Vue-routers, Vuex-store) for seamless typing

[2] Presentation on Apache Hadoop Goes Realtime at Facebook (with team of 4)

- Analysis Facebook's paper *Apache Hadoop Goes Realtime at Facebook*; discussed reasons behind Facebook's decision to move to Hadoop and HBase, and methods of testing and optimization HBase based on their needs
- Presented a 30 mins demo about the result in course with professors and classmates questions.

[3] Review & Presentation on Cloud Database Management System (DBMS) (with team of 8)

- Comparatively analyzed main Cloud DBMS products (AWS, Azure, and Google Cloud) regarding data warehousing, database migration, time series database, in-memory data store, and big data query as a service
- Conducted case studies on cloud-based DBMS users with Alibaba & Taobao + Google & Schrodinger
- Broke down status quo and future development of cloud DBMS market and cloud technology research

LEADERSHIP & ACTIVITIES

Journalism/ Newspaper Club

Sep 2015 - Jun 2018

Vice Chair Journalist and Editor

Beijing, CN

- Spearheaded organizational innovation, focused on improving the efficiency of the conduct interviews includes celebrity interviewer; contribute articles; edit the newspaper, school WeChat official account & website
- Won 5-Star Club (Best) in high school

Member of Big Data & Analytics Association; Member of ACM club

Medical Insurance Development, Amcare Women's & Children's Hospital

Jul 2017 - Feb 2018

Research Assistant

Beijing, CN

- Conducted market research, analysis & orientation; Discussed product design concept; made product models; learned professional finance knowledge on medical insurance

Regional Games for the Disabled

Reginal Games for the Disabled

Aug 2017 - Sep 2017

Volunteer Organizer

Beijing, CN

STEAM Course at Tsinghua University, 3D Print Concentration

Sep 2015 - Feb 2016

Participant

Beijing, CN

- Attended 3D print course; learned the scientific theories & calculation of FDM; Drew 3D printer components by software; learned 3D printer installation & debugging

SKILLS, CERTIFICATIONS & OTHER HIGHLIGHTS

Programming: Java, Python, C,

Tools: SQL, MATLAB, Git, Linux

Certifications: Deep Learning Specialization (by deeplearning.ai on Coursera), IBM Data Science (by IBM)

- Languages: English (Fluent), Chinese Mandarin (Native)