Jimmy Zhang

Email: jimmyzhang2003@gmail.com
Mobile: +1-718-675-0506
Portfolio: jimmyzhang.me
GitHub: github.com/jimmyzhang2003
LinkedIn: linkedin.com/in/jimmy-zhang03

Education

- **Columbia University in the City of New York** — New York, New York
  Bachelor of Arts, Double Major in Computer Science and Cognitive Science — *Sep 2021 - May 2025*
  - **GPA**: 4.07/4.00 (Dean's List)
  - **Relevant Coursework**: Advanced Programming, Data Structures in Java, Computational Linear Algebra, Discrete Mathematics, Calculus III
  - **Activities and Societies**: Application Development Initiative (Committee), Columbia Undergraduate Science Journal (Editor), Columbia Science Review (Editor), Intramural Volleyball

Experience

- **Memorial Sloan Kettering Cancer Center** — New York, New York
  Software Engineering Intern — *June 2022 - Present*
  - Build a command-line interface, a scalable plugin system, and a plugin generator to automate access to MSK resources and code generation for developers across the entire Digital, Informatics & Technology Solutions (DigITs) division.
  - Work alongside software engineers in the Polaris team to design, build, and deploy a clinician-facing Connected Care Dashboard application, with a focus on front-end development, UI design, and user authentication.

- **The Data Science Institute at Columbia University** — New York, New York
  Data Science Institute (DSI) Scholar — *Sep 2021 - Present*
  - Selected by Dr. Daniel Freedberg to first-author a research study identifying associations between specific antibiotic classes and risk for *C. difficile* infection, wielding a MarketScan database of over 40 million unique patient records.
  - Utilize SAS software to preprocess, visualize, and analyze over 200 million antibiotic prescriptions from inpatient, outpatient, facility header, and pharmaceutical claims data spanning 12 years.

- **Icahn School of Medicine at Mount Sinai** — New York, New York
  Research Assistant — *Dec 2019 - Present*
  - Spearheaded a team of 6 MDs, research scientists, and engineers in Professor Kuan-lin Huang's Computational Omics Lab and implemented 4 different machine learning algorithms to predict COVID-19 diagnosis, culminating in a first-author publication in Nature Scientific Reports.
  - Leverage over 150,000 rows of longitudinal data taken from the largest Alzheimer's database in the US to predict mortality in dementia patients and identify the key clinical features associated with mortality across 8 dementia subtypes.

- **AI Camp** — Palo Alto, California (Remote)
  Software Engineering Intern — *Jul 2021 - Dec 2021*
  - Oversaw 5 members in AI Camp's inaugural Talent Incubator cohort to create a Gmail Add-On that filters college emails.
  - Designed a web scraper to fetch data of 4,454 colleges from a college search website, parsed the data into a SQLite database, and implemented sorting logic that reduced email filtering time by over 40%.
  - Wrote and published a Medium article with 60+ claps on behalf of the cohort detailing the internship experience.

Projects

- **MSK CLI** A command-line interface for internal use by developers at Memorial Sloan Kettering Cancer Center, reducing overhead time spent in the browser and while writing redundant boilerplate code. **Tech**: Node.js, Typescript, oclif.
- **HTTP Web Server** An HTTP web server built from scratch, employing socket programming to serve static HTML pages and dynamic pages queried from a database lookup server. Re-implemented as an Apache module. **Tech**: C, Apache, Linux.
- **College Email De-Spamifier** A Gmail Add-On that revamps high school students' inboxes, filtering college emails based on user-specified criteria (acceptance rate, tuition, etc.). **Tech**: Python (Django, Beautiful Soup), Google Apps Script, SQLite.
- **Personal Portfolio** A snapshot of my work, goals, and personal life. **Tech**: React.js, Next.js, Tailwind CSS.

Publications

- **Zhang J**, Jun T, Frank J, Nirenberg S, Kovatch P, Huang KL. Prediction of individual COVID-19 diagnosis using baseline demographics and lab data. *Sci Rep* 11, 13913 (2021). https://doi.org/10.1038/s41598-021-93126-7
- **Zhang J**, Chen L, Gomez-Simmonds A, Yin MT, Freedberg DE. Antibiotic Class-Specific Risk for Community-Acquired Clostridioides difficile Infection in the United States from 2008 to 2020. *In Submission.*
- **Zhang J**, Song L, Huang, KL. Predictive Models and Features of Patient Mortality across Dementia Types. *In Submission.*

Skills

- **Languages**: **(Fluent)** English; **(Working Proficiency)** Mandarin Chinese, Spanish
- **Technical Skills**: research, machine learning, full-stack software engineering, data science, data visualization
- **Technical Languages**: Python, Javascript, Typescript, R, Java, HTML/CSS, C, SAS, LaTeX, SQL
- **Frameworks and Technologies**: Node.js, React.js, Next.js, Tailwind CSS, Bootstrap, oclif, scikit-learn, pandas, tensorflow, matplotlib, numpy, tidyverse (dplyr, ggplot2, etc.), caret, mice, SAS macro
- **Tools**: Git, Github, Docker, Postman, Vim, Unix, Visual Studio Code, RStudio, SAS Studio, Overleaf