


Yihan Shi

✉ yihan.shiwork@gmail.com |  linkedin.com/in/yihan-shi |  yihan-shi.github.io

Education

Duke University

Durham, NC | 2019 – 2023

B.S. in Statistics (High Distinction), Minor in Computer Science & Neuroscience | GPA: 3.93/4.00, Cum Laude

- **Relevant Coursework:** Data Structures & Algorithms, Computer Systems, Machine Learning, Artificial Intelligence, Advanced Data Visualization, Statistical Inference, Probability, Innovation & Entrepreneurship, Education Anthropology

Skills & Languages

Programming languages: Python, Java, C, C#, ASM, R, JavaScript, HTML/CSS; **Skills:** Jenkins CI/CD, Git, React, MS Office, Figma, Jira, Canva, AWS, Tableau, \LaTeX ; **Languages:** English, Mandarin Chinese, Japanese; **Database:** SQL;

Work Experience

Mastercard Data & Services | Software Engineer

Arlington, VA | Aug 2023 – Aug 2024

- Developed full-stack web application features using the .NET Framework for enterprise customers such as McDonald's, Walmart, and Bank of America, enabling large-scale document permission management, lookup, editing, uploading, and deletion services
- Revitalized user experience by refactoring CSS, HTML, and JavaScript code to align with industry-leading accessibility and UX standards
- Led an innovative research project using variational autoencoders (a type of artificial neural network) to predict counterfactual campaign outcomes as part of a company AI initiative

Carta Healthcare | Data Science Intern

San Mateo, CA | May 2022 – Aug 2022

- Developed a data visualization dashboard with 10 progress-tracking metrics and presented at a client-facing Product Demo session
- Communicated core needs with the engineering team to speed up their data querying process by designing SQLAlchemy wrappers to scalably query OLAP
- Designed a template and workflow document for 2 NLP scientists to auto-generate HTML webpages for machine learning pre-training
- Reduced 30% of the time bioinformaticians spent on housekeeping tasks by implementing a comprehensive and robust data cleaning process to reduce platform test errors systematically
- Saved 50% of the manual labor of patient profile simulation by building and maintaining an infrastructure that streamlined and scaled 100x simulation based on real-world distribution

Duke University, Statistical Science Dept. | Research & Teaching Assistant

Durham, NC | Aug 2021 – Aug 2023

- Authored technical tutorial for doctors on recurrent event survival analysis methods to examine large claim datasets
- Authored manuscript [in review] on the impact of confounding medical interventions in electronic health records data

Leadership and Extracurriculars

Duke Technology Scholars Program, Duke University

Vice President & Scholar

Jan 2021 – Jan 2023

- Managed 5 tech interest groups (i.e., Tech for Social Good, Data Science) with 170+ students by organizing 4 new member education sessions with other campus technology clubs and assigning monthly deliverables to subgroup leaders
- Recruited 9 subgroup leaders and mentored them on planning events, creating social media campaigns, and tracking goals effectively using Trello
- Oversaw 8 events in 2022, such as food for thought and technical workshops, by setting event agendas for subgroup leaders and creating outreach worksheets

Technical Project

DukeHub Student Portal 3.0

Sep 2021 – Dec 2021

- Created an R Shiny app that improved user experience of picking classes compared to the current student portal DukeHub 2.0 by identifying user needs iteratively through campus-wide surveys and designing a product roadmap
- Provided insights to 2408 Duke courses through interactive data visualization features including enrollment cap comparison, diversity of the subject areas, and expected traveling route and time on weekdays
- Collaborated with 2 student developers on designing and implementing front-end interfaces (i.e., type boxes, drop-down lists) and connected the back-end server to Google Maps API