

ALAN LIU WU

Livingston, NJ | +1 862-485-9127 | alanlwu@seas.upenn.edu | [linkedin.com/in/alanlwu](https://www.linkedin.com/in/alanlwu) | github.com/alan-w25 | alan-wu.me

EDUCATION

University of Pennsylvania

Master of Science in Engineering in Data Science

Activities: Resident Assistant

Philadelphia, PA

Expected May 2026

Rutgers University- New Brunswick

Bachelor's of Science in Computer Science and Data Science, minor in Statistics

Honors: Summa Cum Laude, Paul Robeson Scholar, Phi Beta Kappa

New Brunswick, NJ

Sep. 2021 – May 2024

TECHNICAL SKILLS

Languages: JavaScript, Python, Java, C, C#, R, SQL, HTML, CSS

Libraries/Frameworks: React.js, Next.js, Node.js, Tailwind.css, Django, Pytorch, Pandas, Numpy, Matplotlib

Developer Tools: Git, Github, MySQL, PostgreSQL, Docker, Azure CLI, Firebase, Expo CLI

EXPERIENCE

Graduate Research Assistant

Wharton School of Business

July 2024 – Sept 2024

Philadelphia, PA

- Scraped and inserted over 12 million records of google review data into PostgreSQL database
- Developed **Tableau** dashboard to visualize data relationships between text and image reviews
- Conducted multi-modal sentiment analysis and visualization techniques to extract insights from 3.5 million restaurant reviews with python and OpenAI API

Undergraduate Research Assistant

Rutgers Cancer Institute of New Jersey

Aug 2023 – May 2024

New Brunswick, NJ

- Analyzed viability of new DNA sequencing technique for lung cancer prognosis to model patient survival times and outcomes with survival support vector machines (ssvm), and random survival forests on 108 cancer patients
- Built heat maps, cluster diagrams, and survival curves to measure statistical significance with **Python**, **matplotlib**, **seaborn**. Hypothesis tested features with log-rank tests and Kaplan-Meier curves
- Tuned L2 strength, learning rate, and kernel of ssvm model with stratified k-fold cv, yielding 0.83 C-Index

Software Engineer Co-Op

ASICS Digital

Jul 2023 – Dec 2023

Boston, MA

- Automated previously manual process to verify metadata overlap with **Python** scripts to inform key design decisions during migration to a new Customer Data Platform. Verified over 2 million records
- Designed an A/B testing framework for email delivery using **Postman API**, enabling the marketing team to analyze engagement patterns; insights gained were utilized to optimize delivery strategies for 50,000+ recipients.
- Authored and optimized **SQL** queries, assisting in GDPR-user compliant deletion for over 1.7 million customers
- Directed 3 team workshops to elevate team skill sets in **Postman API**, **SFMC integrations**, **Git VCS**

PROJECTS

LangTrain | *React Native, Expo CLI, TailwindCSS, OpenAI API*

- Released an AI-powered Mandarin Chinese language teaching mobile application with 50+ AI-generated lesson plans based on user topic and difficulty preferences using **OpenAI API** and **Google TTS API**
- Managed image, audio, and lesson object data through CRUD operations through firebase using **Firebase API**

Chinese MNIST Digit Detection App | *Next.js, Fastapi, Pytorch, Docker, Azure*

- Deployed full stack web application to accomplish classification of 15 Chinese hand-written digits
- Built custom CNN architecture that achieved validation accuracy of 0.92
- Containerized app with **docker** compose and used **Azure** web services and container instances for deployment

Online Retail Analysis | *Python, Pandas, Numpy, Matplotlib, Sci-kit Learn*

- Conducted data pre-processing on UCI online retail dataset to ensure quality and consistency, implementing RFM (Recency, Frequency, Monetary) analysis to segment 4,000 customers based on purchase behavior
- Developed a **Random Forest** predictive model, fine-tuned with grid-search cv, achieving a 0.76 R-squared value

Clinical Heart Failure Bayesian Analysis | *R, RMarkdown, rstanarm, ggplot2*

- Executed bayesian logistic regression techniques, implementing an MCMC simulation-based model to predict patient survival from a dataset of 299 heart failure cases
- Achieved 0.87 accuracy using Bayesian logistic regression model, incorporating prior knowledge

INTERESTS AND ACTIVITIES

Hobbies: Cooking, Ultimate Frisbee, Reading, Golf, Learning New Languages, Running, Hackathons