LILLIAN YU

github.com/leeleeian | linkedin.com/in/lillian-yu-cmu 408 - 674 - 7829 | lyu2@andrew.cmu.edu

EDUCATION

Carnegie Mellon University || *Pittsburgh, PA* – **GPA 4.0** *August 2023 - Grad. May 2026*

- **Major:** B.S. in Computer Science, **Minor:** Computational Finance, Dean's List Fall 2023, Spring 2024
- **Programming Coursework:** Data Structures and Algorithms, Computer Systems, Algorithm Design, Imperative Programming in C
- **Mathematical Coursework:** Probability in Computing, Discrete Math, Linear Algebra, Differential Equations, Computational Finance

EXPERIENCE

Full-Stack Software Intern || *RentVest (Startup)* *May 2024 - Sept 2024*

- Implemented a real-time embedded payment system using Stripe API and direct ACH with MongoDB database collections.
- Accelerated property rendering and transaction time by 12% by restructuring the user-to-property data pipeline.
- Established new backend endpoints and set up webhooks for real-time refund and purchase events.
- Conducted web scraping of 50,000 properties in the US using querying, proxy servers and caching techniques.
- Led pre-seed funding rounds for data subscription service; secured 2 big data vendor firms.

TA for 15122 Programming in C || *Carnegie Mellon University* *Jan 2024 - May 2024*

- Led ~20 students in programming labs each week; hosted weekly office hours, helping 10-20 students each shift.
- Improved content for final exam and updated assignment content and rubric; graded weekly assignments.

Data Analysis Intern || *UC Berkeley Materials Science Lab* *June 2023 - August 2023*

- Broadened data integration by compiling 40,000 new chemicals and computed their combined thermonized reaction cascades.
- Developed and visualized enthalpy coefficients using weighted linear combinations in matplotlib.
- Updated and cross-validated data retrieval of over 50 chemical groups from REST API and Materials Project API.

Machine Learning and Data Science Intern || *BU RISE Internship* *July 2022 - August 2022*

- Improved the lab's outlier detection accuracy by 45.55% using Isolation Forest and Support Vector Machine algorithms.
- Enhanced data retrieval and classification efficiency tenfold by developing a data collection process using Pandas and Naive Bayes.
- Achieved 88.52% classification accuracy by employing KNNs, demonstrating generalizability to larger patient cohorts.
- Performed multivariate regression and survival analysis on over 70 parameters, yielding statistically significant results.

Computer Vision Researcher || *UCSB Summer Research Academies* *June 2020 - July 2020*

- Enhanced license plate recognition by increasing edge detection accuracy by 23% via a repurposed Multi-task Cascaded Convolutional Networks (MTCNNs) with improved feature extraction and transformation layers.
- Reduced plate identification processing time by 14% on large datasets, even under challenging lighting.
- Implemented real-time parallel processing by training CNN embeddings using Single Shot MultiBox Detectors (SSD).
- Coauthored and published research findings in the IJHSR journal.

SKILLS

Computer Languages	*[Proficient]* Python; Java; C; R *[Experienced]* C++; Javascript; HTML/CSS; Typescript; MATLAB
Frameworks & Libraries	React JS; NumPy; Seaborn; PyTorch; Flask; Express.js; Matplotlib
Developer Tools	Linux; Git; Conda; VSCode; Visual Studio; Eclipse; Jupyter Notebooks; Google Colab; Vim

ADDITIONAL EXPERIENCE AND AWARDS

HackCMU ACM@CMU: Best Use of MongoDB Atlas *September 2023*

- Created a website in a team for VC firms to connect with space startups, using MongoDB Atlas, HTML/CSS, and JS.

Junior Developer for ScottyLabs Marketplace MVP *September 2023 - December 2023*

- Developed a Marketplace app for CMU students; deployed shopping cart feature by creating RESTful API endpoints for cart operations.