Gary Lin

San Diego, CA, USA

✉ g5lin@ucsd.edu   ⬡ Atlae   in Gary-J-Lin   ⓘD 0009-0003-1150-3578

Experience

Data Science Intern   *Remote*

MICROSOFT CORPORATION, INC. (VIA CENTIFIC), CENTRAL FRAUD AND ABUSE RISK   June 2024 - September 2024

- Developed and implemented retrieval-augmented generation (RAG), finetuning, and head-to-head Elo evaluation pipelines with cutting-edge models such as Phi-3, LLaMa, Mistral, and GPT-4o.
- Integrated in-house AI solutions over their current Copilot model, reducing API fees by ~100% and increasing quality by 15%.
- Found sufficient conditions for small language models to outperform large language models that can be 10 or 100 times bigger in specialized tasks, will appear in an upcoming issue of the Microsoft Journal of Applied Research.

Undergraduate Student Researcher   La Jolla, CA

UC SAN DIEGO, EDWIN SOLARES BIOINFORMATICS + AI (ESB.AI) LAB   April 2024 - Present

- Researching machine learning applications in bioinformatics, focusing on training LLMs to identify mutations in VCF files.
- Labeled over 500 images in LabelStudio annotating segmentation masks and bounding boxes of avocado trees for CV.

Undergraduate Teaching Assistant   La Jolla, CA

UC SAN DIEGO, DEPARTMENT OF MATHEMATICS   September 2023 - Present

- Taught calculus to 2-4 sections (each with up to 40 students) per academic quarter, also proctored and graded exams.
- Developed interactive learning materials for students' understanding and provided constructive feedback and active listening with students with a wide variety of math ability, resulting in a consistent >90% rating from student evaluations each quarter.

Projects

AspectiBot   ⬡ Onteia/AspectiBot

DEVELOPER/MAINTAINER   September 2022 - Present

- Rewrote 70% of, then began maintaining, a bespoke open-source Discord and Twitch chat bot using Java with Maven.
- Introduced version control to the project with GitHub, wrote 100% of the documentation, and created a ChatGPT integration.

Pricebook Automation Project   Closed Source

DEVELOPER/ON-CALL MAINTAINER   August 2022 - Present

- Automated a workflow using Python that converted Excel flat files into formatted LaTeX tables in a 100+ page price book.
- Saved the company 80% labor time from manually editing sales and price numbers from Excel spreadsheets.

REWARD (Raising Eastern World Assembly Recognition and Development)   (Closed Source) ⬡ The-East-Pacific

LEAD PROJECT ARCHITECT   June 2021 - May 2024

- Led and managed about 5 technical and coding projects for The East Pacific, a roleplaying community with over 5,000 members.
- Used Python to scrape the NationStates API and automate a process that saved 90% labor time and increased returns by 200%.

Skills

Languages	**Python**, **Java**, C#, C/C++, JavaScript, TypeScript, HTML, CSS, Bash, SQL, Rust, Golang
Libraries	NumPy, Pandas, Matplotlib, PyTorch, Keras, OpenCV, BeautifulSoup, Requests, Maven
Programs	Visual Studio Code, Git, GitHub, AWS, Google Cloud, Microsoft Azure, Linux

Education

University of California, San Diego (UCSD)   La Jolla, CA

B.S. IN MATH-COMPUTER SCIENCE & BIOLOGY WITH A SPECIALIZATION IN BIOINFORMATICS   September 2021 - June 2025 *(Expected)*

Relevant Coursework (GPA 3.7/4.0):

Data Structures, Algorithm Design, Machine Learning, Artificial Intelligence: Probabilistic Reasoning, Deep Learning, Recommender Systems, Systems Programming, Server-Side Technologies, Online Database Analytics, Web Client Languages, Software Tools, Principles of Programming Languages *(graduate)*, Advanced Compiler Construction *(graduate)*, Linear Algebra, Probability, Statistics, Real Analysis *(graduate)*, Abstract Algebra, Numerical Analysis, Graph Theory, Combinatorics.