# JAMES GERONIMO

🜙 +1 (925) 470-7668 🛛 jegeronimo@berkeley.edu 🔚 linkedin.com/in/james-geronimo 🞧 github.com/jegeronimo

### EDUCATION

#### University of California. Berkelev

Bachelor's in Computer Science, Bachelor's in Data Science

Relevant Coursework: Data Science, Machine Learning, Artificial Intelligence, Probability Theory, Linear Algebra, Data Analytics, Data Structures and Algorithms, Efficient Algorithms, Intractable Problems, Machine Structures, Computer Security, CS Education Honors: Dean's List (College of Letters & Science), Honors to Date, Upsilon Pi Epsilon (Computer Science Honors Society) Organizations: Data Science Society (Vice President), Computer Science Mentors (Senior Mentor), Open Project, Intramural Sports

#### SKILLS

Languages: Python, SQL, R, Java, JavaScript, C, C++, Golang, Scheme, Ruby, x86, RISC-V, HTML, CSS, LATEX Data Science/Machine Learning: Pandas, NumPy, SciPy, RegEx, Matplotlib, Seaborn, Plotly, Scikit-Learn, PyTorch, TensorFlow, LangChain, Hugging Face, NLTK, GENSIM, TextBlob, SpaCy, OpenCV, Pix2Text, Tesseract-OCR, PyMuPDF/Fitz, Power BI, Tableau Software Development: Git, React.js, Tailwind CSS, Jekyll, Flask, Heroku, Microsoft Azure, JUnit, Selenium, Playwright, OAuth 2.0

### EXPERIENCE

#### Amgen June 2025 – August 2025 Incoming Data Science Intern Remote University of California, San Francisco Feb 2025 – Present Research Assistant Remote • Evaluate large language models (LLMs) robustness via conversational red-teaming and multi-turn jailbreaking techniques • Implement an agentic framework to automate biomedical task evaluations and enhance LLM robustness under the Abbasi Lab Dec 2024 – Present AmigoAI Machine Learning Intern Berkeley, CA • Devise and build an automated end-to-end testing framework using Playwright to detect regressions and improve test coverage • Engineer a robust and dynamic web scraper that mimics real-user behavior by emulating user agents, preloading consent cookies, and bypassing JavaScript-rendered content through direct HTML data extraction using Scrapy, BeautifulSoup, and Selenium • Scrape 1000+ data entries into JSON files for fine-tuning a custom-built LLM for standardized exam question generation • Construct a cross-format file processing workflow to extract text from files using PvMuPDF/Fitz and Pix2Text • Architect a computer vision pipeline that tracks real-time user progress to contextualize query processes in a fine-tuned LLM UC Berkeley College of Computing, Data Science, and Society Jun 2024 – Present Undergraduate Student Instructor (uGSI) Berkeley, CA • Teach 1200+ students in the upper-division course Data 100 C, covering data processing, statistical/probabilistic foundations, exploratory data analysis/visualization, feature engineering, dimensionality reduction, predictive modeling, and optimization • Streamline course website & navigation by developing in Jekyll and automating deployment workflows through GitHub Pages • Lead discussion sections to 40+ students, host office hours, debug DataHub issues, and maintain general course infrastructure UC Berkeley Data Science Undergraduate Studies Feb 2024 – Present Berkeley, CA Software Developer • Partner with four El Camino College professors to create educational data science modules teaching probability distributions, data visualization, classification, statistical modeling, and business analytics under the Data Science Modules C team • Design and construct two data science modules deployed through Jupyter for the undergraduate interdisciplinary course LS 22 🗗 • Introduce 500+ students every semester to fundamental programming practices, statistics, and data-driven decision-making Universal Health Services May 2024 – Aug 2024 Remote Data Scientist • Achieved <18.5% and <25% average error in financial forecasting by deploying AutoRegressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) neural network models, optimized with extensive EDA and fine-tuning • Delivered Pyramid Analytics integration, Python script, and walkthrough video; leveraged mentorship from C-Suite executives PROJECTS Oct 2024 - Dec 2024 **Comprehensive News Chatbot** $\square$ | Python; HuggingFace, GENSIM, LangChain, SpaCy • Built an interactive text generation tool to generate coherent, accurate summaries of news articles from the past decade

- Extracted keywords from queries and leveraged Sentence-BERT to retrieve sentence embeddings for document chunking
- Applied Retrieval-Augmented Generation (RAG) with cosine-similar articles to contextualize GPT-3.5 Turbo through LangChain

## GCN and LLM Directory 2 | Python, JavaScript; PyTorch, React, Flask, Heroku

- Created an AI-assisted directory for 100+ members with a Python/Flask backend, React frontend, and Heroku deployment
- Employed a Graph Convolutional Network (GCN), achieving a 0.87 ROC AUC score for link classification accuracy
- Incorporated Google OAuth 2.0, fetched from an AirTable database, and generated contextualized responses from GPT-3.5 Turbo

Feb 2024 – Apr 2024

May 2026 GPA: 3.872/4.0