Brandon Yan

510-398-9092 | branyan@seas.upenn.edu | LinkedIn | GitHub | Portfolio | Fremont, CA | US Citizen

Education

University of Pennsylvania Philadelphia, PA

Bachelor of Science in Engineering in Computer Science *May 2027*

- GPA: 3.82/4.00, Courses: Data Structures and Algorithms, Discrete Math, Big Data Analytics (**TA**), Algorithms, Computer Systems, Machine Learning, Databases, Operating Systems, Artificial Intelligence, Machine Perception
- Awards: USA Computing Olympiad Gold(556/6,024), 2x AIME(620/24,817), ACSL Finalist(99/1,247)

Experience

Software Development Engineering Intern June 2025 - August 2025

Amazon — Artificial General Intelligence Team *Sunnyvale, CA*

- Led integration of Titan architecture into Nova models, extending 128k context window to **over 2 million** tokens
- Productionized an adapter layer for Nova's distributed pipeline (multi-node A100), enabling long-context runs
- Created novel test-time learning inference framework, providing **16.8× accuracy** increase at long contexts
- Enhanced memory & batching to deliver **15%** faster inference and **20%** lower GPU use at Nova scale

Vice President, Penn Spark (Advanced Projects Team) Jan 2025 – Present

University of Pennsylvania *Philadelphia, PA*

- Direct 50+ developers and designers across technical projects, workshops, and mentorship programs
- Work directly with teams to integrate APIs and optimize pipelines across full-stack and ML projects

Undergraduate Research Assistant May 2024 – Present

University of Pennsylvania *Philadelphia, PA*

- Train 3D U-Net models on CBCT scans for tooth segmentation, improving accuracy of surgical planning tools
- Collaborate with clinicians to refine medical labels, improving segmentation reliability for surgical use
- Present work on bullet fragment detection at ORS 2025 Meeting, highlighting medical contributions
- Engineer web scraping tool using Python and Selenium to extract financial ratios for over **10,000** companies

Software Engineer for AI Training Data May 2024 - May 2025

Scale AI *San Francisco, CA*

- Optimized LLM production by interpreting over 100 model responses for the Code Extensions and Ostrich teams
- Utilized APIs like Google Maps, Google Flights, and YouTube, enhancing tool to handle **50+** diverse query types
- Performed quality assurance on training datasets by fixing edge cases, ensuring robustness in LLM applications

Software Engineer Intern July 2021 - August 2021

Cisco Systems *San Jose, CA*

- Designed an application with two Cisco employees to enable startups to discover untapped markets globally
- Assembled density heatmap using Google Maps API, HTML, CSS, and JS, visualizing data for **1,000+** locations
- Integrated NodeJS and GoLang to process 10,000+ IP addresses, leveraging AWS, EC2, and S3 for web scalability

Projects

Spotify Popularity Prediction | *Python, Pandas, Scikit-Learn, XGBoost* November 2024 – Present

- Built a machine learning pipeline to predict song popularity by preprocessing a dataset of 100,000+ Spotify tracks
- Developed Random Forest and XGBoost models, optimizing prediction accuracy through hyperparameter tuning

Pop-Plan | *NextJS, OpenAI, LangChain, Pinecone, Clerk, Firestore, Whisper, JavaScript* June 2024 – October 2024

- Developed and integrated a chatbot feature (Poppy) that personalizes itineraries, increasing user interaction by **30%**, by utilizing retrieval-augmented generation (RAG) to pull data from user profiles, amassing **500+ users**

Gesture Genie | *Python, Flask, Git, BeautifulSoup, HumeAPI, MediaPipe, OpenCV* June 2024 – July 2024

- Trained a neural network and data models using TensorFlow, enhancing gesture recognition accuracy
- Integrated OpenCV for image processing and HumeAPI for emotion recognition, cutting processing time

Technical Skills

Languages: C++, Java, Python, C, Golang, R, HTML, Javascript, CSS, SQL, Typescript
Frameworks: React, Node.js, TailwindCSS, Flask, JUnit, Material-UI, FastAPI, NextJS
Developer Tools: Git, Docker, Pinecone, Amazon Web Services, VS Code, IntelliJ, Eclipse, Google Firebase
Libraries: pandas, NumPy, Matplotlib, OpenCV, LangChain, scikit-learn, BeautifulSoup, Selenium