

☎ (669)-291-0140

📍 Cupertino, CA

EDUCATION

The Johns Hopkins University

B.S. Computer Science, B.S. Applied Mathematics & Statistics (Dean's List)

• **Focus Areas:** Software Engineering, Natural Language Processing, Statistics and Statistical Learning

• **Teaching (OOP – C/C++):** Guided 350+ students; Assessed 300+ assignments; Scripted 5+ autograding tests; Provided OH debugging support

PROFESSIONAL EXPERIENCE

Palantir | Gotham

Incoming Software Engineering Intern (Fall 2024)

Intel | Linux 3D Graphics Performance

Open-Source Software Engineering Intern

• Engineered 5+ functional testing tools (screen record/capture, render diff, etc.) for 3D graphics benchmarking and optimization using Python (OpenCV, FFmpeg, PIL, MP) and C++ (Vulkan, OpenGL, Kernel Interface) with Mesa Graphics drivers, increasing testing coverage by 47+%

• Designed a synchronized data collection system mapping frames with I/O events, creating a 10k+ input database for training gameplay scripts

• Contributed to a semi-supervised ML system using RLHF and transfer learning with PyTorch, performing feature engineering and data augmentation to automate gameplay script generation, reducing time by 70%, increasing accuracy by 30%, and expanding test suite by 10+

• Integrated 2 profiling tools into the functional suite enhancing shader profiling techniques (runtime analysis, shader execution, etc.) using RenderDoc and Perf, expanding API record/replay for improved performance analysis, optimizing the graphics rendering pipeline by 15+%

Scale AI | Impala (Outlier)

Artificial Intelligence Model Trainer

• Crafted 100+ high-quality training datasets to enhance sophisticated AI models to accurately understand and process complex math concepts

• Evaluated and ranked model responses for accuracy, relevance, and factuality using statistical scripts, boosting model capabilities by 12+%

Google | Computer Science Research Mentorship Program

Research Scholar

• Gained insight into the R&D cycle, emphasizing on CI/CD pipelines, code QA, and the reliability/performance of multi-distributed ML systems

• Engaged in 10+ panels on computing research, delving into system architecture design, optimization, and scalable ML software solutions

StudyFind | Internal Software Development

Junior Software Developer (Full-Stack)

• Refined 3+ RESTful microservice architectures for real-time distributed data processing, increasing transaction throughput accuracy by 18%

• Architected 4 server-side internal tools using Node.js with Express, employing A/B testing to refine database interactions and workflow

• Led code reviews on backend reliability, resolved 7+ critical Firebase Realtime Database API issues, and optimized Jenkins CI/CD pipelines

SoKat | jArvis AI Equity Research Platform

Machine Learning Engineer Intern

• Coded 5+ Azure RESTful APIs with Nginx load balancing/Circuit Breaker fault tolerance, ensuring instant financial forecasting for 2K+ firms

• Developed 3+ robust internal tools for automated financial data scraping and aggregation from 10+ sources (SEC filings, 8K, etc.) using Python (BeautifulSoup, Scrapy), employing SQL indexing, NoSQL sharding, and query caching/batching techniques to boost backend speed by 18+%

• Boosted backend throughput by 35+% using Redis in-memory caching, asynchronous tasks, and exponential backoff to enhance reliability

• Researched and fine-tuned 20+ financial sentiment analysis, summarization, and QA transformer models, applying ML metrics to embed top-performing models into existing backend architecture with Docker and Kubernetes, thereby improving accuracy and analytical scope by 18+%

RESEARCH EXPERIENCE

Johns Hopkins University – Center for Language and Speech Processing

Undergraduate Researcher

• Implemented multi-GPU processing across distributed cluster environments using Slurm scripts and CUDA, leveraging parallel computing to train/test/validate transformer-based neural networks in PyTorch and generate 10+ model checkpoints for NLP information retrieval tasks

• Programmed a scalable NLP pipeline using Tevatron/Elasticsearch, yielding a 20%+ ETL performance boost, to process a large-scale dataset of 80M+ web articles from 50+ languages that facilitates the development of a cross-lingual information retrieval model for citation generation

• Crafted a human-inspected 2k+ passage dataset from using HuggingFace for instruction-following LLM tasks and benchmark experiments

FEATURED PROJECTS

MedTheia – Visualizing Dense Clinical Texts

• Engineered a stable-diffusion text-to-image extension prototype using PyTorch fine-tuned on 15K+ clinical texts and images from EHRs

• Devised a back-end data processing pipeline using NLP libraries (NLTK, SpaCy), MySQL, and 5+ open-source libraries (Apache, Airflow, etc.)

GastroCrawl – Food Crawl Planner

• Utilized AWS EC2 for scalable hosting and SageMaker to deploy ML-driven route optimization, integrating C++ backend for high-speed tasks

• Constructed a Node.js backend using AWS Lambda for serverless route calculation, MongoDB for efficient data storage, and 20+ RESTful APIs

ChatJH – JHU Questions, Answered

• Created a lightweight BeautifulSoup/Scrapy-based web scraper mining 3T+ words from 5+ JHU subdomains, generating 700K+ QA-pairs

• Fine-tuned a GPT-3.5 chatbot for JHU-related queries, refactored code reducing latency by 35%+, and integrated 3+ RESTful APIs for access

TECHNICAL SKILLS

• **Languages:** Python, Java, C++, HTML, CSS, JavaScript, C, MATLAB, Bash, SQL, TypeScript, Assembly

• **Selected Libraries & Frameworks:** BeautifulSoup, Scrapy, PyTorch, Django, Flask, NLTK, tkinter, Requests, PyTest, JUnit, Apache, STL, Bootstrap, PyData (NumPy, Pandas, scikit-learn, etc.), MERN (MongoDB, Express.js, React.js, Node.js), Next.js

• **Database & Cloud:** MySQL, PostgreSQL, Firebase, SQLite, Redis, Azure, AWS, Google Cloud Platform (GCP), Heroku

• **Tools & Platforms:** CUDA, Docker, Slurm, Kubernetes, Postman, Git, Jira, Databricks, Airflow, Jenkins, Terraform, Ansible

Benjamin Chang

benjaminchang.dev • bchang26@jhu.edu

in benjaminchang7

[benchang323](https://github.com/benchang323)

Baltimore, MD

Expected 05/25

Washington, DC

06/24 – Present

Hillsboro, OR

05/24 – Present

San Francisco, CA

01/24 – 05/24

Mountain View, CA

09/23 – 12/23

New York, NY

03/23 – 09/23

Woodstock, MD

06/23 – 08/23