

Augustin Tsang

+1 (510) 484 1734 ✉ augustintsang@gmail.com 📄 [/augustin-tsang](https://www.linkedin.com/in/augustin-tsang) 🌐 [/augustintsang](https://github.com/augustintsang) 🌐 [augustin.codes](https://www.augustin.codes)

Education

University of California, Berkeley

Computer Science, B.A. | GPA: 3.765

August 2021 – PRESENT

Berkeley, California

Skills and Relevant Coursework

Languages, Frameworks and Tools: Python, PyTorch, TensorFlow, Java, Golang, C/C++, Node.js, JavaScript, React, SQL, AWS, JUnit, Linux Kernels

Relevant Coursework: Database Systems, Data Structures and Algorithms, Efficient Algorithms and Intractable Problems, Operating Systems, Machine Learning, Computer Security, Computer Architecture and Machine Structures

Experience

Deepr

Software Engineer

May 2024 – PRESENT

Berkeley, CA

- Developed an advanced LLM-based song matching system using Sentence Transformers, achieving a 25% improvement in matching accuracy and significantly enhancing the UX for Deepr's music recommendation platform.
- Built and optimized scalable data pipelines with AWS, reducing data ingestion time by 30% and ensuring efficient, real-time processing of high data volumes.

EpiNu Project, the Blum Centre for Developing Economies

Software Engineer

January 2024 – May 2024

Berkeley, CA

- Spearheaded the development of a prenatal micronutrient security algorithm using a computer vision model in the DRC under Dr Sonia Navani, that pinpoints which micronutrients are lacking and their geographical region.
- Benchmarked and integrated a high-performance distributed database, building a comprehensive micronutrient database on SQLite, and integrating it with a custom YOLO model in PyTorch, achieving a 55% reduction in computational load for effective use in low-bandwidth scenarios.
- Utilized a K-Nearest Neighbors algorithm to offer alternative meal suggestions that optimize nutritional intake.

Climformatics

Software Engineer

October 2023 – December 2023

Berkeley, CA

- Trained and tested ML models evaluating the ice cover sensitivity given changes in albedo perturbation and cloud cover, improving model prediction accuracy by 15%, which was presented at the American Geophysical Union conference.
- Conducted timeseries analysis on spatial data to produce effectiveness maps of albedo perturbation using a 1D model across Arctic grid points, optimising locations for artificial albedo enhancement, cutting implementation costs by 20%.
- Processed ensemble data with xarray and used Sabalcore supercomputing over the cloud to compare linear regression, decision trees, and random forest approaches to increase prediction accuracy using python and sklearn.

Projects

Secure File Storage and Sharing System | Golang, RSA, JSON, Datastore, Keystore APIs

August 2024

- Implemented user authentication, file storage, retrieval, sharing, and access revocation.
- Ensured data confidentiality and integrity using RSA encryption, digital signatures, and authenticated encryption.
- Prevented file tampering, swapping, and user structure corruption with robust error handling and data structures.

EEG Focus for YouTube | Python, Pandas, Muse SDK, OpenAI API

March 2024

- Placed 1st at Neurotech@Berkeley/NVIDIA hackathon by utilising the Muse S EEG Headset to monitor EEG signals, processing the signals into users' attention levels.
- Synchronised attention levels with YouTube timestamps and captions, and automatically generates simplified summaries for complex content during periods where brainwaves are unfocused.

CNN Model to Assess Severity of Breast Cancer from Biopsy Images | Python, scikit-learn

December 2023

- Spearheaded the creation of a Convolutional Neural Network model, using TensorFlow and the ResNet-50 architecture, to assess the severity of breast cancer from biopsy images. Implemented on a dataset of 175,000 images to predict high-risk cases, facilitating better clinical decision-making.
- Achieved a diagnostic accuracy of 90.8% in validation tests, significantly outperforming traditional diagnostic methods. The project showcased potential for substantial improvements in patient treatment plans and clinical outcomes, verified using a subset of 500 biopsy slides from the Providence Cancer Institute.

RookieDB Database System and Optimisations | Java

October 2023

- Built database implementation project with a focus on sequential transaction execution.
- Core optimisations included B+ tree indices, query optimization, multigranularity locking for concurrent transaction execution, and database recovery.