

# Sergio Zavala

[linkedin.com/in/sergiozavala1](https://www.linkedin.com/in/sergiozavala1) | [szaval314@gmail.com](mailto:szaval314@gmail.com) | Soledad, CA | <https://github.com/sezavala> | [sergiozavala.tech](https://sergiozavala.tech)

## EDUCATION

### California State University Monterey Bay | Seaside, CA

**Expected: May 2026**  
**GPA: 3.78**

B.S. in Computer Science  
Data Science Concentration

- **Relevant Courses:** Data Structures and Algorithms, Computer Architecture, Database Systems, Operating Systems, Internet Programming, Intro to Data Science, Advanced Machine Learning, Software Engineering
- **Research Interests:** AI for Science, Scientific Machine Learning, Machine Learning for Security, Computer Vision, ML Systems, High-Dimensional Data Modeling

## RESEARCH

### Capstone Undergrad Research | California State University, Monterey Bay

**April 2026 - Present**

- Built an NLP pipeline to analyze and summarize course evaluations, extracting actionable insights from noisy, unstructured feedback
- Evaluated summarization and sentiment models across datasets (CSUMB, UCLA), designing metrics to measure performance and generalization
- Developed a benchmarking framework to compare model outputs and assess insight quality across diverse evaluation formats

### NIH NHGRI GREAT Undergrad Research | University of California, Santa Cruz

**June 2025 - Present**

- Developing a scalable CLI tool to project genomic annotations onto pangenome graphs across 462 human assemblies.
- Building a Python pipeline to map millions of annotations and benchmark projection accuracy, reducing linear-reference bias.

## POSTER PRESENTATIONS

**Sergio Zavala, Prajna Hebbar, Dr. Benedict Paten.** "Pangenome Annotation Toolkit: A solution for mapping gene annotations through a graph."

- Presented at NHGRI Training Meeting, 2026 (St. Louis, MO)
- Presented at SACNAS NDiSTEM, 2025 (Columbus, OH)
- Presented at CSUMB Summer Research Symposium, 2025 (Seaside, CA)

## PROFESSIONAL EXPERIENCE

### AgRobotics | Data Collection Engineer

**March 2025 - Present**

- Incoming Data Collection Engineer, working with the deployment of autonomous field robots and LiDAR-based sensing systems to collect high-resolution agricultural data for analysis and machine learning applications.

### Computing Talent Initiative (CTI) | Software Engineer

**February 2025 - Present**

- Engineering CTI-sys data management system; implemented JavaScript sanitization to prevent malicious requests and built endpoints using Pandas and returned formatted information onto Sheets for analytics.

### California State University, Monterey Bay | Instructional Student Assistant

**August 2025 - Present**

- Supported Data Structures and Computer Architecture courses through office hours, labs, grading, and Slack support.
- Developed a CLI tool to visualize binary conversions and arithmetic; authored AI usage guidelines for coursework.

### EBSCO Information Services | CodeDay, Software Engineer Intern

**October 2024 - November 2024**

- Built a Python + Pandoc pipeline to convert arXiv data to JSON for Longformer training; deployed Flask API, cutting reading time by 80%.

### Oppia | CodeDay, Software Engineer Intern

**June 2024 - August 2024**

- Resolved critical resource-loading bug impacting 1M+ users and implemented TypeScript validation checks and unit tests.

### Open Energy Dashboard | CodeDay, Software Engineer Intern

**January 2024 - February 2024**

- Authored integration tests verifying PostgreSQL-API consistency in Dockerized staging; streamlined energy-data pipeline used by 10k+ users.

## PROJECTS

### GaelCraves | Java, TypeScript, Python, PostgreSQL, Docker

**November 2025 - December 2025**

- Engineered a production-grade Spring Boot backend featuring JWT security, PostgreSQL integration via Supabase, and Docker containerization, delivering a scalable REST API deployed to Heroku during a CodeDay internship.

### CAPTCHA Robustness Analysis | TensorFlow, Keras, VGG16, OpenCV

**December 2025**

- Built a multi-output CNN (VGG16 backbone) to decode noisy CAPTCHA; improved accuracy 20%+ via targeted augmentation and transfer learning.

**AMPLIE** | React Native, Node.js, Vector Databases**October 2025**

- Developed a full-stack, emotion-aware music platform using Groq and Chroma DB to perform semantic track retrieval with 85% match accuracy, integrating multi-agent orchestration via Fetch.ai to synchronize group playlists in real-time.

**Personal Website** | *EXPRESS.JS, MYSQL, EJS***May 2025 - June 2025**

- Developed a full-stack portfolio with a JWT-protected CMS enabling in-browser blog and project management.

**Gen AI Feedback Tool** | *GOOGLE APPS SCRIPT***February 2025 - June 2025**

- Built an AI-driven grading pipeline in Google Apps Script, reducing feedback turnaround from weeks to ~10 minutes for 100+ students.

**PROFESSIONAL DEVELOPMENT, VOLUNTEER, & MENTORSHIP**

---

**ACM SACMAT 2026** | Artifact Evaluation Committee Member**April 2026 – Present**

- Selected to evaluate research artifacts for an ACM security conference; assessed reproducibility, correctness, and documentation of submitted systems.

**United Way of Monterey County** | Volunteer**Spring 2026 - Present**

- Developing a data tracking and analytics system to monitor participation, outcomes, and website-integrated visual summaries.

**NIH NHGRI GREAT Scholars** | Mentor**Spring 2026 - Present****Mentors in Tech Program** | Mentee**Fall 2025 - Present****Computing Talent Initiative (CTI Accelerate Program)** | Student Assistant**Summer 2025 - Present****CS++ Program** | Student Ambassador**Fall 2022 - Fall 2023****HONORS & AWARDS**

---

**NSF ExLENT CTCtoAI Student Fellowship****2025 - Present****Outstanding Research Poster - CSUMB Summer Symposium****2025****JFW Cares Scholarship****2025 - 2026****NIH NHGRI GREAT Scholar****2024 - Present****CSUMB College of Science Dean's List Recipient****2022 - Present****SKILLS**

---

**Languages & Tools:** Python, Java, JavaScript/TypeScript, SQL, Bash, Git, Docker**Frameworks & Libraries:** FastAPI, Flask, Django, Spring Boot, React, Angular**Databases:** PostgreSQL, MySQL, SQLite**Data & ML:** Pandas, NumPy, scikit-learn, matplotlib, Tensorflow (Keras), Jupyter Notebook**Software Engineering:** Full-stack development, REST APIs, Auth, Testing (Mocha/Chai), CI/CD