

# LAWRENCE WONG

U.S. Citizen

✉ [lawrence.wong.2004@gmail.com](mailto:lawrence.wong.2004@gmail.com) ☎ (832) 366-6226 🔗 [linkedin.com/in/lawrencewongg](https://www.linkedin.com/in/lawrencewongg)

🐙 [github.com/Laroldz](https://github.com/Laroldz) 🌐 [lawrencewongg.net](https://lawrencewongg.net)

## Education

### Texas A&M University

*Master of Science in Computer Science*

Aug. 2026 – Dec. 2027 (expected)

*College Station, Texas*

- Graduate Coursework GPA: 4.00/4.00

### Texas A&M University

*Bachelor of Science in Computer Science*

Aug. 2022 – May 2026

*College Station, Texas*

- GPA & Awards: 3.93/4.00, Summa Cum Laude, Undergraduate Research Scholar

**Relevant Coursework:** Artificial Intelligence, Computational Data Science, Cloud Computing, Deep Learning, Machine Learning, Natural Language Processing, Parallel Computing

## Publications

- **SHIELDS: Multi-Agent OS Security Hardening.** [arXiv preprint](#), 2026.
- **Analyzing Sensor-Driven Activity Recognition in Swimming Schools Using Deep Learning.** Undergraduate thesis, 2026.

## Experience

### Systems and Technology for Medicine and IoT (STMI) Lab

*Graduate Researcher*

Aug. 2026 – Present

*College Station, Texas*

- \* Built the preprocessing pipeline for MIMIC-IV EHR data, filtering and formatting ~**158M** raw lab events across ~**93K** ICU stays into model-ready sequences on TAMU HPRC Grace.
- \* Developing **transformer-based models** to forecast **cardiac biomarker** trajectories for heart failure risk assessment.
- \* Performed exploratory analysis on clinical EHR data, visualizing lab-event frequency and coverage to identify dataset limitations and gaps in existing approaches.

### OptiSigns Digital Signage

*Data Engineering & Applied AI Intern*

May 2025 – Aug. 2025

*Houston, Texas*

- \* Built an LLM-based industry classification pipeline using the OpenAI API, achieving ~**85%** classification accuracy through prompt iteration and systematic evaluation using **Braintrust**.
- \* Engineered a scheduled BigQuery inference pipeline on DigitalOcean using a **cron job**, integrating web-scraped and internal SQL data to classify ~**200** customers daily.
- \* Delivered enriched industry classifications for ~**30,000** historical and new customers to BigQuery, enabling downstream analytics and customer segmentation at scale.

### Sketch Recognition Lab

*Student Researcher*

Aug. 2024 – May 2026

*College Station, Texas*

- \* Engineered sliding-window statistical and frequency-domain features from resampled wearable accelerometer/gyroscope data, building the feature pipeline behind downstream activity classification.
- \* Implemented a hierarchical multistage human activity recognition pipeline in **PyTorch** for land/water detection and swimming activity classification, achieving an **88.9% macro-F1** on the land-vs.-water stage.
- \* Presented research to judges during a Student Research Week event and wrote an undergraduate thesis for the Undergraduate Research Scholars program.

### Texas A&M University

*Peer Teacher*

Aug. 2023 – May 2026

*College Station, Texas*

- \* Delivered foundational Python instruction to more than **90** first-year engineering students, created quizzes, and coordinated grading of assignments, quizzes, and exams to strengthen core programming concepts.
- \* Led large-scale test reviews averaging **60** participants and produced supplemental content to reinforce lecture materials.
- \* Moderated the largest TAMU Engineering Discord server, supporting a **5,500-member** community of students outside of class hours and advising on curriculum and major-selection (ETAM) questions.

## Projects

### SHIELDS: Multi-Agent OS Security Hardening | *Python, OpenSCAP, LLMs, Linux, SSH, XML* May 2026

- \* Built three Linux environments with increasing **OpenSCAP STIG** vulnerability levels to benchmark **LLM-based** security remediation across models and workloads.
- \* Implemented a multi-agent remediation workflow that parsed OpenSCAP findings, triaged vulnerabilities, generated and executed fixes, validated system safety, and iteratively rescanned failed findings using feedback from prior attempts.
- \* Benchmarked multiple **LLMs** with detailed logging of model outputs and remediation behavior, and contributed to an arXiv paper analyzing how model behavior, tool use, and iterative feedback affect operating-system security hardening.

## Weakly Supervised ICD-Codable Clinical Text Detection | *PyTorch, Transformers, NLP* May 2026

- \* Developed a weakly supervised clinical NLP pipeline to identify ICD-codable segments in long medical notes using document-level labels and gated-attention multiple instance learning, and co-authored to an IEEE-formatted paper evaluating the approach for extracting clinically important information from medical notes.
- \* Implemented text preprocessing, engineered to a TransICD-based initial model, and tuned model hyperparameters for training on long clinical documents segmented into fixed-length token windows.
- \* Trained a frozen **BioClinicalBERT** encoder with a two-layer MLP classifier, achieving **89.9% AUROC**, **77.2% F1**, and **81.5% accuracy** on a labeled test set.

## Pneumonia and Breast Cancer Image Classification | *PyTorch, LibAUC, NumPy* May 2024

- \* Leveraged **LibAUC** to attain **96%** and **85%** AUC for breast cancer and pneumonia classification in the MedMNist dataset, writing a comprehensive research paper detailing methods, experiments, results, and conclusions.
- \* Optimized hyperparameters by testing various optimizers (e.g **momentum**, **learning rate**) using GridSearch and experimenting with image augmentations (e.g **rotation**, **scaling**, **contrast**) to improve both models by  $\sim 4\%$ .

## NotOnMyBingoSheet | *React, Node.js, Express.js, MySQL, AWS Cognito, IAM & RDS* Jan. 2025

- \* Engineered a web application to create, save, and share Bingo sheets.
- \* Integrated **AWS Cognito** for secure user authentication, ensuring confidential access and sharing permissions.
- \* Architected a full-stack solution using **AWS RDS** to store user information and bingo sheets in multiple SQL tables and deployed on **Render**.

## Technical Skills

---

**Languages:** Python, SQL, C++, Java, JavaScript, TypeScript, R, HTML/CSS, LaTeX

**Machine Learning & Data:** PyTorch, Hugging Face Transformers, scikit-learn, NumPy, Pandas, OpenAI API, Braintrust, LLM Prompting, Jupyter

**Cloud & Infrastructure:** AWS, BigQuery, Docker, DigitalOcean, Linux, SLURM, GitHub

**Web & Databases:** React, Node.js, Express.js, MySQL, PostgreSQL