My (Chiffon) Nguyen

San Francisco, CA, USA | hi@mychiffonn.com | github.com/mychiffonn | mychiffonn.com

Research Interests

Human–AI systems that reflect and support diverse human patterns: **multilingual and multicultural AI**, **AI alignment** (pluralistic alignment, cooperative AI, scalable oversight, evaluation, mechanistic interpretability), **socially-aware and responsible AI** (human-AI collaboration, human-AI interaction ML fairness).

Education

Minerva University, College of Computational Sciences Sep 2021 — May 2025

B.S in Computational Sciences (Machine Learning and Statistics), GPA: 3.7/4.0 *San Francisco, CA, USA*

- Relevant Coursework: Machine Learning (A), AI Ethics, Bayesian Modeling (A), Statistical Modeling and Causal Inference (A), Optimization Methods (A), Probability and Statistics (A-), Software Engineering
- Certifications: Natural Language Specialization (deeplearning.ai, 2024), Machine Learning Specialization (2022)
- Self-study: AI Alignment (ARENA, 2025), Introduction to AI Alignment (Bluedot Impact, 2025)

Research Experience

AI Research Fellow Sep 2025 — Present

Algoverse AI Research *Remote*

- Researching AI alignment under Eyon Jang, aimed for a ICML 2026 workshop paper

Machine Learning Research Assistant Jun 2024 — Aug 2024

AI & Mixed Reality Lab, Landshut University of Applied Sciences *Landshut, Bavaria, Germany*

- Implemented 3D object detection pipeline using LiDAR point clouds and PointPillars algorithm in PyTorch and NVIDIA TAO Toolkit for autonomous navigation research
- Conducted comparative analysis of model performance across standard vs. synthetic datasets

Causal Inference Research Intern (Replication and Extension) Nov 2023 — Dec 2023

Minerva University (Advisor: Professor Alexis Diamond) *Remote*

- Replicated and extended Chrisinger (2021)'s causal inference of Philadelphia's beverage excise tax effects on SNAP benefit redemption using synthetic control in R, analyzing policy impacts across 4+ counties and 50+ months of longitudinal data
- Identified critical limitations in dataset reliability and **magnitude discrepancies** between original and replicated results
- Conducted new leave-one-out robustness analysis on synthetic control models, showing model instability

Teaching Experience

Minerva University

Lead Tutor, PR51 Programming with Python Spring 2025

- Taught weekly labs for 40+ first-year students of diverse learning backgrounds, covering computer systems, Python fundamentals, object-oriented programming, debugging, security, and data structures
- Extracted 20 data-driven pedagogical insights using Google Drive API, Google Sheet trackers, student and tutor surveys, improving hands-on learning and student engagement for the next class iteration

Lead Teaching Assistant, FA50/FA51 Critical, Statistical and Algorithmic Thinking Fall 2023 - Spring 2024

- Guided 150+ students annually in formal logic, probability and statistics, algorithmic thinking, and simulation, through office hours, personalized feedback on 25 quizzes; assisted professors in grading 6 assignments annually.

Community Engagement & Services

- Aug 2025 - Present: **Communication and Design Engineer** at SEACrowd (AI research for Southeast Asia), building website and managing crowdsourced data annotation for underrepresented languages
- Feb 2025: Dataset contributor for Humanity's Last Exam benchmark dataset and SEA-VL

Selected Projects

Astro Academic Portfolio Theme for Researchers (github.com/mychiffonn/website) Aug 2025 — Oct 2025

High-performance academic theme, enabling fast, accessible, and multilingual publishing of publications and technical blogs while ensuring top-tier SEO (80+) and Lighthouse (100) scores using Astro, TailwindCSS and shadcn/ui

Mnemonic Generation for Vocabulary Learning (github.com/mychiffonn/mnemonic-gen) Oct 2024 — Mar 2025

- Designed an AI chatbot that generated mnemonic devices for learning and memorizing vocabulary, synthesizing 50+ papers across linguistics, psycholinguistics, language education, and large language models
- Applied **chain-of-thought distillation pipeline**: generated 10k synthetic reasoning examples from teacher model DeepSeekR1, fine-tuned student model Gemma3-1b using LoRA via trl and unsloth libraries
- Implemented **Direct Preference Optimization** on 500 annotated preference pairs, achieving statistically significant improvements in alignment with learner judgments on cognitive science principles such as memorability and imageability

SeizureSavvy (github.com/mychiffonn/SeisureSavvy) Feb 2024 — Apr 2024

- Managed a team of 4 to build a Progressive Web App for intuitive seizure management with machine learning-based predictive alerts in Flask (Python), React, and Chakra UI
- Enhanced data logging and medication tracking, increasing user data accuracy by 35%
- Conducted code reviews in Python and React JavaScript, cutting 40% critical bugs and enhancing application stability

Technical Skills

- **Programming Languages:** Python, TypeScript, R, SQL, Bash
- **Machine Learning & Statistics:** PyTorch, trl, unsloth, TransformerLens, scikit-learn, LangGraph, LlamaIndex, SciPy, PyMC
- **Web:** React, Astro, Express, FastAPI, PostgreSQL, Flask, Jekyll, Tailwind CSS, shadcn/ui
- **Tools & Technologies:** Git, Docker, Render, Railway, Netlify, LaTeX, Zotero, Typst