

EDUCATION

University of California, San Diego (UC San Diego) *September 2020 – June 2024*
La Jolla, CA
Bachelor of Science in Data Science, Minor in Linguistics

Stanford University *September 2024 – Present*
Palo Alto, CA
Master of Science in Computer Science. Anticipated Graduation June 2026.

RESEARCH

Faithfulness in Chain of Thought Reasoning *May 2023 – May 2024*
Laboratory for Emerging Intelligence. UC San Diego. La Jolla, CA
Principal Investigators: Dr. Leon Bergen and Dr. Ramamohan Paturi

- First author on this project, presented our work at the First Conference on Language Modeling.
- Designed and executed experiments, leveraging cognitive methodology to analyze unfaithfulness in LLM-generated reasoning. Findings demonstrated a clear dissociation between faithful and unfaithful behaviors.
- Developed a taxonomy of LLM reasoning errors and recovery behaviors, enabling a previously unachievable level of detailed analysis.
- Managed a team of annotators to apply this taxonomy for efficient data analysis and performed statistical analysis to identify the significance and effect size of experimental results.

Robust Noisy Chain of Thought *October 2023 – December 2023*
Laboratory for Emerging Intelligence. UC San Diego. La Jolla, CA
Principal Investigators: Dr. Leon Bergen and Dr. Taylor Berg-Kirkpatrick

- Co-led a team to improve the robustness of Large Language Model (LLM) reasoning in noisy environments, building on my previous findings about hallucination behavior in reasoning.
- Conducted literature reviews and integrated findings into project development.
- Implemented and executed Python-based exploratory experiments, fine-tuning pre-trained LLMs to target the effect of text dropout in training on test-time reasoning ability.

Cognitive Biases in Syntax Production *June 2019 – September 2019*
Language and Cognitive Neuroscience Lab. UW Madison. Madison, WI
Principal Investigators: Dr. Maryellen Macdonald and Dr. Mark Seidenberg

- Assisted with the development of 4 experiments which drew connections between vision, language processing, and motor control systems in humans.
- Utilized natural language corpora to design experimental stimuli for an experiment measuring response time to nonstandard syntactic constructions
- Created proof-of-concept experimental design demonstrations, using Python and Microsoft PowerPoint.
- Performed data cleaning, annotation, and analysis using Microsoft Excel and Praat, an audio-speech processing software.

PUBLICATIONS

- **Evelyn Yee**, Alice Li, Chenyu Tang, Yeon Ho Jung, Ramamohan Paturi, and Leon Bergen. *Dissociation of Faithful and Unfaithful Reasoning in LLMs*. In First Conference on Language Modeling, 8 Oct. 2024. url: <https://openreview.net/forum?id=IPZ28ZqD4I>

PRESENTATIONS

- **Evelyn Yee**. *Chain of Thought & Verbalized Machine Reasoning*. University of California – San Diego: Laboratory for Emerging Intelligence Reading Seminar, 4 Oct. 2023.
- **Evelyn Yee**. *The Role of Reasoning and Truthfulness in Large Language Models*. University of California – San Diego: Summer Research Conference, 17 Aug. 2023.

TECHNICAL SKILLS

Programming Languages – Python, Java, C++, R, JavaScript, HTML, CSS, SQL, DAX
Software Tools – Microsoft Excel, Microsoft Power BI, Workato, Snowflake, Matillion, Praat

TEACHING EXPERIENCE

Course Assistant

September 2024 – Present

Stanford University

Course: Introduction to Probability, ACE (MS&E 120ACE)

Instructor: Dr. Ellen Vitercik (Assistant Professor)

- Serve the Additional Courses for Engineers (ACE) program, providing supplemental instruction for undergraduates from marginalized and underserved backgrounds.
- Design and present course materials, including weekly lectures and additional tutoring as needed to improve student outcomes.

Undergraduate Tutor

July 2023 – August 2023

UC San Diego

Course: Data Structures & Algorithms for Data Science (DSC 30)

Instructor: Kevin Lin (Visiting Assistant Professor)

- Facilitated discussion sections, held office hours, and graded student assignments.
- Received unanimously positive final evaluations from students and the instructor.

OTHER EMPLOYMENT

Data Analytics Intern

June 2022 – March 2023

The Boldt Company, Appleton, WI

- Created dynamic reports in Power BI to analyze budget usage, employee demographics, and project flow.
- Developed, deployed, and wrote documentation for two automated data archiving processes using Workato and Snowflake.
- Ran audits on project technology to ensure consistency across company data.
- Self-organized and coordinated with other IT members to deliver results according to internal customer deadlines.

Webmaster

April 2020 – Ongoing

Randolph Community Clinic, Randolph, WI

- Manage digital infrastructure for the clinic’s telehealth and COVID-19 vaccination efforts.
 - Perform regular website administration to increase functionality and web traffic for patients.
 - Collaborate with the clinic management team to develop digital solutions for business and healthcare issues.
-

SERVICE

Peer Mentor

January 2023 – Ongoing

Data Science Student Society, UC San Diego

- Advise new students through their transition to college and the data science major.
 - Organize regular meetings to teach essential skills like goal setting and time management.
-

RELEVANT COURSEWORK

Math (Linear Algebra, Vector Calculus & Analytical Geometry), **Probability & Statistics, Data Structures & Algorithm Design, Machine Learning** (Probabilistic Modeling, Recommender Systems, Deep Learning, Natural Language Processing), **Signal Processing**

COURSE PROJECTS

Context-Aware Plagiarism Advisory Tool

Fall 2024

Trustworthy Machine Learning (CS 329T). Stanford University

<https://github.com/evelynyee/plagiarism-rag>

- Identify potential plagiarism source documents, analyze textual relationships to detect plagiarism, and prepare an informational report to advise instructor decisions.
- Implement a custom pipeline for plagiarism detection in Python, using retrieval augmented generation (RAG) and chain of thought.
- Develop evaluation criteria for context relevance, groundedness, answer relevance, and confidence, without need for ground-truth labels.

Localized Noise Function for Stable Diffusion

Winter 2022

Deep Learning for Natural Language Understanding (LIGN 167). UC San Diego

<https://github.com/evelynyee/lign-167-final-project>

- Explore the effect of custom noise functions, to promote minimally-destructive inpainting functionality in text-to-image diffusion models.
- Integrate LLMs and computer vision systems into a pipeline for image processing using the diffusion model framework.

Etymology Visualization Dashboard

Spring 2023

Introduction to Data Visualization (DSC 106). UC San Diego

<https://evelynyee.github.io/dsc106-final>

- Design and build an interactive dashboard to allow open-ended exploration and analysis of a large, graph-based etymology dataset.
- Write custom HTML and CSS and perform live data processing using D3.js.
- Leverage GeoJSON data to place linguistic information in a real-life context.