

Oishani Bandopadhyay

obandopadhyay@ucsd.edu | github.com/oishani-b | [linkedin](#)

Education

University of California, San Diego	Jun 2025 - Present
Masters of Science in Computational Social Science	
University of California, San Diego	Jun 2021 - Jun 2025
Bachelor of Arts in Linguistics	
Bachelor of Science in Cognitive Science with a Specialization in Machine Learning and Neural Computation (Double Major)	
<ul style="list-style-type: none">• Graduated with Distinction• Linguistics Honors, Sophomore Honors, Provost Honors (x6 quarters)• Student Employee of the Year Runner-Up Award 2025	

Research Interests

Natural Language Processing, Computational Linguistics, Phonetics, Automatic Speech Processing, Multilingual Speech, Indian Dialects, Sociolinguistics

Research Experience

Honors Thesis: Evaluating ASR Systems Accuracy Using Word Error Rate Across Indian and Standard US Dialects

advised by Dr. Will Styler

- Designed research experiment based on phonetic differences in Indian English and mainstream US dialects, distributing a survey collecting over 1,800 audio recordings from participants
- Analyzed differences in transcription accuracy using WER (word error rate) to evaluate 4 popular ASR (automatic speech recognition) systems across over 40 prompts with varying levels of complication and dialectal variation
- Used statistical modeling to find effects of prompt and language background differences on lower accuracy, finding lower accuracy on systems for Indian English speakers

Evaluating Talker Similarity Using ASR for Multitalker Accent Adaptation

advised by Dr. Sarah Creel

- Building Whisper ASR pipeline for automated transcriptions and using WER with human judgements of speaker intelligibility to create a framework evaluating talker similarity

- Using Praat and neural forced alignment techniques to find phoneme boundaries in talker speech, at single word and sentence level, across varied speaker language backgrounds
- Collecting speech from over 40 L1 and L2 English speakers with Mandarin language backgrounds, transcribing and evaluating speech to create golden dataset

Graduate Capstone: Multilingual Keyword Search

advised by Dr. Will Styler

- Developing open multilingual (KWS) keyword search dataset, adapting existing ASR (automatic speech recognition) datasets to evaluate KWS on a multilingual benchmark
- Using techniques in ASR including finite state transducers to get phoneme-word transcriptions, MFCC (mel frequency cepstral coefficients) representations of speech, HMMs (hidden Markov models) to model speech, DTW (dynamic time warping) for signal similarity for KWS

Developing GPT-Scientist Library for Qualitative Coding in Social Sciences Research

advised by Dr. Nadia Polikarpova

- Investigating gaps in trustworthiness, reliability, effectiveness, and scalability of LLM-based tools for qualitative coding tasks in social sciences research
- Identifying both open and closed coding tasks in qualitative research across various domains, benchmarks for evaluating LLMs as coders, existing frameworks on inter-rater reliability, and areas for improvement in developing automated qualitative coding systems

Understanding Everyday Technology Practices of Adults with ADHD: Insights for Assistive

Technology Design

advised by Dr. Munyaka

- Contributed to literature review on existing assistive technologies for adults with ADHD, centering the experiences of neurodivergent technology users and proposing an approach to developing inclusive and effective design frameworks
- Used mixed methods approach including LDA (Latent Dirichlet Allocation) topic modeling and qualitative analysis to analyze over 96,000 Reddit threads from communities of individuals with ADHD, finding areas of improvement in assistive technologies

Teaching

Instructional Assistant

Sept 2024- Jan 2025

COGS 108, Data Science in Practice, UCSD

- Worked closely with Professor Jason Fleischer to facilitate COGS 108 - Data Science in Practice for 770 students, leading weekly discussion sections with programming exercises and review of core concepts in data science and machine learning

- Improved quality of end-to-end student projects, advising 26 student groups by providing feedback on checkpoints, troubleshooting data cleaning, analysis, model selection, report writing in weekly office hour sessions
- Assisted with organizing AFA accommodations for students throughout the quarter, responding to student questions on discussion boards

Lead Undergraduate Writing Consultant

Aug 2023- Jun 2025

Teaching & Learning Commons, UCSD

- Worked with over 200 students in one-on-one tutoring sessions, advising students on a wide range of writing projects, from research and technical writing to analytical and persuasive essays
- Provided mentorship to 17 new tutors and held office hours to support cohort of over 40
- Interviewed and evaluated candidates in hiring process as part of 2024-25 Student Staff Panel
- Launched weekly Writing Strategy Sessions initiative, presenting interactive workshops on practices in college level writing to groups of over 30 students, hosted community writing hours
- Assisted in writing proposal for and planning a round table discussion on working with multilingual students, addressing issues with grammar and standard American academic writing conventions for the SoCal Writing Centers Association Conference 2025

Work Experience

AI Agent Developer Intern, NeuralSeek

Aug 2025 - Sept 2025

- Won 1st place in 200 interns in final demonstration, presenting chat interface that analyzed over 50 pages of California AI bills & reports to provide regulatory recommendations
- Built multi-agent workflow to automate legal research, increasing efficiency of compliance analysis with minimal hallucination, achieving 94% response accuracy on semantic scoring

Machine Learning Engineer Intern, Quantiphi

Jun 2024 - Oct 2024

- Developed financial document understanding tool extracting data from over 30 page financial reports with tables, charts, and text, improving interpretation accuracy for documents by 10%
- Implemented multi-modal neural networks to process domain-specific documents, integrating Google Cloud & Vertex AI, using vector database and custom chunking RAG pipeline

Generative AI Research Intern, Third Eye Advisory

Jun 2023 - Sept 2023

- Prepared research reports on Gen AI techniques in collaboration with Senior Consultants
- Researched latest developments in multimodal AI applications, LLMs, image processing

Data Science Intern, Third Eye Advisory

Jul 2022 - Sept 2022

- Analyzed large company survey data using Excel and Python, visualizing and identifying trends

- Found global trends in adoption of NLP and automation, regulatory effects, and productivity
-

Presentations & Workshops

Finding your Niche in Research - Cognitive Science Student Association Winter Talk Series 2025

Discussed research opportunities at UCSD within Cognitive Science, Linguistics, and related labs or departments for undergraduate students, presented latest progress on Honors Research

ChatGPT Workshop - Leveraging Generative AI - Cognitive Science Student Association Workshops

Designed an interactive workshop for students at all levels of experience with generative AI tools to showcase best practices in prompt engineering and interesting capabilities of ChatGPT 3.5

The Grammar Conversation - Writing Hub Tutor Education Session

Showed findings from a year-long pedagogical study on grammar conventions and difficulties for students from different educational backgrounds with Standard American Writing

CVs, Resumes, and Cover Letters - Black Resource Center Workshop, with the Writing Hub

Led a workshop on best practices for writing graduate school application materials, and writing resumes and cover letters for job applications, working with students in an interactive follow-up session

Leadership & Activities

Co-President, Cognitive Science Student Association

Jun 2023 - Jun 2025

- Led growing group of over 180 students in Cognitive Science and associated departments to host professional events with academic and industry speakers, and community building
- Organized the annual Cognitive Science Spring Conference at UCSD with over 70 attendees, hosted 10 speaker talks on research and industry projects , facilitated 2 round-table discussions
- Hosted Lab Open House with over 150 attendees and 12 research labs across linguistics and cognitive science presenting posters, Winter Talk Series with professors discussing ongoing work
- Created and managed first cohort of interdisciplinary projects with 30 students, providing mentorship, technical feedback, research guidance, and presentation templates

Technical Consultant - Program Manager, DS3 Consulting x Alama AI

Apr 2024 - Oct 2024

- Led team of 12 data science consultants & machine learning developers at Alama AI in building academic advising chatbot based on collected data about over 50 student experiences with academic advising to identify gaps & needs

- Built backend pipeline, utilizing web-scraping, data collection, cloud platform for user-specific chats & NLP techniques to deploy functional chatbot with high response relevance, using context from user profile & university requirements
- Organized weekly sprints with founders & full-stack team, managing technical documentation, report writing & code reviews

Digital Literacy Educator, TechUp

Mar 2024 - Sept 2024

- Organized digital literacy courses for over 40 unhoused members of the San Diego community in partnership with TechUp and the San Diego Rescue Mission
- Certified students in basic technology, running workshops on computers and software tools

Data Visualization Specialist, The Guardian

Oct 2023 - Oct 2024

- Created effective visualizations for data in news articles published by the UCSD Guardian
- Using survey and research data, made relevant plots with Matplotlib, Seaborn, Python, and Microsoft Excel, enhancing reader engagement and understanding

GAP Ambassador

Jun 2023 - Dec 2023

- Mentored international and exchange students entering UCSD for the first time through ISPO (International Students and Programs Office) to familiarize them with campus facilities
 - Answered questions, organized mentorship schedule, and maintained weekly newsletters
-

Skills

Technologies:

Python (Pandas, Numpy, Scikit-Learn, NLTK, PyTorch, HuggingFace, LangChain, Matplotlib, Gensim, SpaCy, PyDub, ffmpeg)

R (RStudio, tidyverse, ggplot, lm)

FindingFive

LaTeX

Praat

Languages:

English

Bengali

Hindi