

Education

2020 - Current **University of California San Diego (UC San Diego), CA.**

PhD, Electrical and Computer Engineering, GPA: 3.96

2025-2026 Dr. William S.C. Chang Best Dissertation Award

Advisor: Prof. Pengtao Xie

2016 - 2020 **Indian Institute of Technology Hyderabad (IIT H), India.**

Bachelor of Technology; Major: Electrical Engineering, Minor: Computer Science and Engineering, GPA: 9.68 (out of 10)

◦ **Academic Excellence Awardee** in academic terms 2016-2017, 2018-2019.

◦ **Second Highest CGPA** in B.Tech Program across all departments (240 students).

Work Experience

Aug 2025–Present **Senior Applied Scientist, Oracle Cloud Infrastructure, AI Science team**, Redwood City, CA, USA.

- Contributed to an **ontology-driven semantic layer** for **agentic business intelligence**, enabling LLM-based data agents to reason over enterprise schemas, business metrics, and ontology constraints across structured and unstructured data.
- Architected and co-led a **proactive insights discovery agent** that synthesizes evidence from enterprise data, logs, documentation, historical repositories, and benchmark datasets to identify recurring patterns, anomalies, and actionable operational insights for engineering support workflows.
- Co-led contributions to **award-winning NL2SQL systems**, achieving **1st place** in the Archer NL2SQL Challenge with **54.96% execution accuracy** and a **6.3-point lead** over 2nd place, and **#1 on Spider 2.0 Lite** with **72.02% execution accuracy**.
- Developed core **NL2SQL planning, critique, and SQL log based verification components** that improved reliable natural-language analytics, reduced hallucinated SQL, and delivered **13.0% average gains** across six benchmarks, with up to **16.7% gains** on ambiguous questions.
- Provided **technical leadership** by defining solution vision, writing research and product proposals, aligning cross-functional stakeholders across research, engineering, and product teams, and driving semantic-layer and agentic BI methods toward product adoption.

Research Internship Experience

June–Sept 2023 **Research Intern, Apple**, Cupertino, CA, USA.

- Pre-trained a **multi-modal LLM** for **on-device deployment**, taking both image and text inputs to improve personalized text predictions, enhancing the smart keyboard's contextual accuracy and overall user experience on **iPhones**.

June–Sept 2022 **Research Intern, Tencent AI Lab**, Bellevue, WA, USA.

- Enhanced **BERT** model performance by **~10%** on **extremely low-resource datasets** through data augmentation, employing a task-dependent similarity matrix, with both this matrix and model weights optimized via a meta-learning-based approach.

May–Aug 2019 **Research Intern, Texas A&M University**, College Station, TX, USA.

Halliburton Engineering Global Program Scholar

- Developed a control algorithm for self-driving cars to navigate crowded environments by leveraging a pedestrian behavior model built with **Inverse Reinforcement Learning**.

Selected Publications

IJCKG, 2025 [OraPlan-SQL: A Planning-Centric Framework for Complex Bilingual NL2SQL Reasoning](#)
Sai Ashish Somayajula*, Marianne Menglin Liu*, Syed Fahad Allam Shah, Sujith Ravi, and Dan Roth

NeurIPS, 2025 [DreamPRM: Domain-Reweightd Process Reward Model for Multimodal Reasoning](#)

Qi Cao, Ruiyi Wang, Ruiyi Zhang, Sai Ashish Somayajula, and Pengtao Xie

EMNLP, 2025 [Improving the Language Understanding Capabilities of Large Language Models Using Reinforcement Learning](#)

Sai Ashish Somayajula*, Bokai Hu*, Xin Pan, Zihan Huang, and Pengtao Xie

EMNLP, 2025 [Can Prompts Rewind Time for LLMs? Evaluating the Effectiveness of Prompted Knowledge Cutoffs](#)

Xin Gao, Ruiyi Zhang, Daniel Du, Saurabh Mahindre, Pengtao Xie, Sai Ashish Somayajula

TMLR, 2025 [TapWeight: Reweighting Pretraining Objectives for Task-Adaptive Pretraining](#)

Ruiyi Zhang, Sai Ashish Somayajula, and Pengtao Xie

- TMLR, 2025 [Downstream Task Guided Masking Learning in Masked Autoencoders Using Multi-Level Optimization](#)
Han Guo, Ramtin Hosseini, Ruiyi Zhang, **Sai Ashish Somayajula**, Ranak Roy Chowdhury, Rajesh K Gupta, and Pengtao Xie
- ICML, 2024 [Token-Specific Watermarking with Enhanced Detectability and Semantic Coherence for Large Language Models](#)
Sai Ashish Somayajula*, Mingjia Huo*, Youwei Liang, Ruisi Zhang, Farinaz Koushanfar, and Pengtao Xie
- TMLR, 2024 [Transformer Architecture Search for Improving Out-of-Domain Generalization in Machine Translation](#)
Yiheng He*, Ruiyi Zhang*, **Sai Ashish Somayajula***, Pengtao Xie
- NAACL, 2024 [Generalizable and Stable Finetuning of Pretrained Language Models on Low-Resource Texts](#)
Sai Ashish Somayajula, Youwei Liang, Abhishek Singh, Li Zhang, and Pengtao Xie
- NAACL, 2024 [AutoLoRA: Automatically Tuning Matrix Ranks in Low-Rank Adaptation Based on Meta Learning](#)
Ruiyi Zhang*, Rushi Qiang*, **Sai Ashish Somayajula**, and Pengtao Xie
- Scientific Reports, Nature, 2024 [Improving Long COVID-Related Text Classification: A Novel End-to-End Domain-Adaptive Paraphrasing Framework](#)
Sai Ashish Somayajula, Onkar Litake, Youwei Liang, Ramtin Hosseini, Shamim Nemati, David O. Wilson, Robert N. Weinreb, Atul Malhotra, and Pengtao Xie
- Scientific Reports, Nature, 2024 [Improving Image Classification of Gastrointestinal Endoscopy Using Curriculum Self-Supervised Learning](#)
Han Guo, **Sai Ashish Somayajula**, Ramtin Hosseini, and Pengtao Xie
- ACL, 2023 [Bi-level Finetuning with Task-dependent Similarity Structure for Low-resource Training](#)
Sai Ashish Somayajula, Lifeng Jin, Linfeng Song, Haitao Mi, and Dong Yu
- TACL, 2022 [A Multi-Level Optimization Framework for End-to-End Text Augmentation](#)
Sai Ashish Somayajula, Linfeng Song, and Pengtao Xie

Invited Talks

- o Natural Language Processing and Vision (NLPV) seminar, University of Exeter
- o UT Austin NLL Reading Group
- o Statistical Visual Computing Group, UC San Diego
- o Indian Institute of Technology, Hyderabad

Teaching, Reviewing, and Service

Teaching Assistantship.

Random Processes, Linear Algebra and Applications, Computational Evolutionary Biology (**Best TA, 10/10 rating**), Statistical Learning I, Deep Generative Models, Linear Electronic Systems.

Reviewer for Top-Tier Conferences.

Conferences: NeurIPS, ICML, ICLR, ACL-ARR, ICASSP.

Student Research Workshop Co-Chair – EACL 2026.

Appointed Co-Chair for the EACL 2026 Student Research Workshop (SRW). Responsible for coordinating the mentorship program, managing submissions and reviews, and supporting author–mentor interactions.

Awards and Honors

- o [OraPlan-SQL](#) wins the [Archer's challenge, 2025](#).
- o Travel grant for ICML 2024, NAACL 2024.
- o Best **Teaching Assistant rating 10/10** for ECE 208, Computational Evolutionary Biology, Spring 2023.
- o Awarded the *Jacob School of Engineering Departmental Fellowship* for PhD studies at UC San Diego 2020.
- o Microsoft Azure Award winner at the [Engineering the Eye-2018 Hackathon](#) 2018.
- o Runners-up in the PwC Challenge on “Smart Cities” at [Megathon](#) 2017.
- o Secured a rank in the top 0.1% among 0.5 million students in the IIT-Joint Entrance Exam 2016.
- o Certificate course in Indian Carnatic music, 2016.
- o Bronze Medal in the National Science Olympiad 2012.

Leadership

- 2025 **Supervision at Dr. Pengtao Xie's Lab, UCSD, UC San Diego.**
Mentoring masters and early year PhD students on various projects, with a strong emphasis on skill development and professional growth.
- 2021 **ENLACE program, UC San Diego.**
Led a team of college and high school students on the project '*Deep Learning Algorithms for Disease Segmentation in Chest X-rays*'; managed data preprocessing, model development, and evaluation, while fostering collaboration and technical skills.
- 2020 **First-Year PhD Representative, ECE Graduate Student Council, UC San Diego.**
Elected to represent and advocate for the needs and concerns of first-year PhD students in the department.
- 2018 **Elektronika Club, Head, IIT H.**
Led the Electronics and Signal Processing Club, fostering innovation and critical thinking in electronics and signal processing among students.
- 2017 **College Counsel, Co-Founder.**
Launched a startup that guided top-ranking students from India's prestigious IIT entrance exam in selecting academic branches, using a personalized approach based on exam rank and personal career aspirations.
- 2018 **Marketing Team, Head, IIT H.**
Led efforts to engage investors in supporting student-led startup initiatives and entrepreneurial ventures.

Skills

- **Machine Learning:** Hugging Face, PyTorch, TensorFlow, spaCy, NLTK, Scikit-Learn, pre-training and fine-tuning *LLMs*, RLHF, PPO, DPO.
- **Programming Languages:** Python, C++, MATLAB, Java, SQL, Bash/Unix, Git.
- **Software Packages:** OpenCV, Jupyter, Pandas, Keras, Numpy, Matplotlib, seaborn.