SACHIT GAUDI

gaudisachit | gaudisac@msu.edu | +1-517-944-7743 | sachit3022.github.io

Professional Summary

Machine Learning Researcher specializing in improving the generalization capabilities of generative models, with a strong background in diffusion models and transformers. Experienced in building and deploying large-scale machine learning systems.

Education

Masters M.S (Thesis) Computer Science	**Michigan State University**, East Lansing, USA	Jan 2023 - Dec 2024
Bachelor of Technology	**Indian Institute of Technology**, Guwahati, India	Jul 2014 - May 2018

Publications Under Review

S. Gaudi, G. Sreekumar, V. Boddeti, "Learning Conditionally Independent Marginals Enables Logical Compositions in Conditional Diffusion Models", **ICLR 2025** [Under Review]

Research Projects

Logical Compositionality in Diffusion models Code and pre-print available upon request

- Proposed the CoInD Algorithm, based on the invariant principle of causality, challenging the conventional conditional independence assumption. Demonstrated that this assumption is frequently violated in practical applications.
- Outperformed existing baselines by a huge margin on unseen compositional generation task.

Adversarial finetuning to mitigate bias in LLM generation , pdf, talk

- Proposed an Adversarial Prompt Tuning Pipeline to reduce bias in LLM generation. This pipeline involves Gumbel-Softmax reparameterization trick, a LoRA adapter, and Stochastic Weight Averaging.
- Achieved a significant reduction in the EOD metric with minimal compromise on generation quality.

Distributed training of LLMs , pdf

- Developed the P-GPT Algorithm in C++ utilizing MPI and OpenMP, implementing a gradient map-reduce parallelization strategy to train large transformer models across multiple servers. P-GPT optimizes parameters for data scheduling, threading, and process parallelism, achieving a parallel efficiency of **95%**.

Novel Domain Generalisation Algorithm , pdf

- Proposed a Novel Domain Invariant Optimization to address spurious correlations, mathematically proving the optimisation transforms into a generalized eigenvalue problem with a closed-form solution. Resulting in an impressive **8%** improvement in generalization accuracy over classic ML algorithms.

NeurIPS 2023 Machine Unlearning Challenge , pdf

- Proposed the Soft-Relabeling Algorithm to effectively remove private information from models. This method achieved a **50%** effectiveness in erasing private data while maintaining a minimal 4% drop in utility.

Professional Experience

Graduate Research Assistant East Lansing, MI

Advisor: Prof. Vishnu Boddeti *Jan 2023 - Present*

- My research can be summarized as "You get what you train for." - Incorrect objective will lead to poor generalization, even when large data is thrown at a problem. Demonstrated that the conditional diffusion objective, without constraints on independence, does not lead to controllability. Additionally, without the right constraints on fairness, models inevitably learn the biases present in the data. Therefore, proposed appropriate objective functions from first principles.

Reliance Jio Haptik Mumbai, India

Machine Learning Engineer - Commerce AI Agent team *Apr 2021 - Dec 2022*

- Spearheaded the Development of an AI Commerce Agent that facilitated chat commerce by seamlessly integrating natural language interactions, eliminating the need for button clicks. Led to a 50% increase in conversion rates. Media
- Built the NLP pipeline featuring a **RAG** (Retrieval Augmented Generation) system. Extracted information from queries, queried ElasticSearch, and generated responses from a knowledge graph.
- Led research and implementation efforts in multi-task BERT-based transformer models for product labeling, named entity recognition (NER), and intent detection in JioMart Search (25M MAUs), resulting in **10%** improvement in **recall**.
- Engineered a collaborative filtering based product personalization module enhancing targeted recommendations, achieving significant improvement with **average first click rank** decreasing from 3.4 to **2.1**.

Software Engineer I *Jan 2020 - Apr 2021*

- Built various microservices, employing async queues for inter-service communication. Orchestrated CI/CD pipelines for software and data delivery. Developed Git actions and wrote bash scripts triggering Jenkins deployments on Azure and AWS. resulting in a **30%** reduction in latency as well as contributing to a **99%** up-time guarantee.
- Developed a solution based on mediator pattern to integrate messages from all chat sources, to a central agent.
- Applied back-track algorithm to enhance Intent detection & utilized Tf-idf for improved semantic extraction.

Smatbot AI Agent Hyderabad, India
Lead Backend Engineer *Jan 2019 - Jan 2020*

- Led architecture and database design for the chatbot. Implemented database sharding, replication, and pushed efforts to make platform No-code, reaching a scale of **10 req/sec**. Contributed to securing a favorable exit for the investors.

Saint Joseph's University Philadelphia, USA
Data Science Summer Research Associate *May 2017 - Jul 2017*

- Built text summarization and sentiment analysis models for hotel reviews, which are now commonly seen on platforms like Amazon and Google. Additionally, developed a predictive model to improve occupancy rates.

Teaching at Michigan State University

CSE 232: Object Oriented programming Spring'24, Fall'23

- Instructed 800 students in C++ language. Developed problem sets and delivered lectures on pointers, classes, IO, strings, objects, and recursion. Course content is available online.

CSE 480: Database Systems Summer'23

- Taught Locks and Object-Oriented Design. Guided students in hands-on creation of SQL database.

Technical Skills

Programming	Python (PyTorch, TensorFlow, sklearn), C++, CUDA, Java, Shell Scripting
Databases & Search	PostgreSQL, MySQL, Redis, MongoDB, Solr, ElasticSearch
Tools & Technologies	Flask, Django, Git, Jenkins, RabbitMQ, Wandb, Jira
Relevant Courses	Computer Vision, Graph Theory, Parallel Computing, Data Structures and Algorithms
Research Notes	Approach to solving constrained objective (pdf). Survey on Bias in LLMs (pdf)

Achievements

Best Employee: Received Big Hand award from Reliance for exemplary quarterly performance, youngest to do so.
Scholarship: Received Scholarship based on Merit from IIT, Kharagpur. (Top 10%).
Joint Entrance Examination: Secured an All India Rank of **2240** out of 1.35 million candidates. (Top 0.17%)

Extra curriculars & Services

Coding: Regularly solve coding challenges on Leetcode (profile), with 500+ problems solved and 5+ unique solutions.
Running: Amateur runner, frequently completing 5K runs with a PB of 30:53, and a time of 1:11:50 for 10K runs.
Technical Blog: I maintain a blog where I explain deep learning concepts like convex optimization and diffusion models with graphical yet intuitive mathematical explanations.
Products I built: Have you thought of how challenging it is to create marketing collateral for your product? Traditionally, it requires an army of people from production to editing. Now, you can create it with just text. That's Surakav.Ai.
Reviewer: NeurIPS 2024, ICLR 2025