Sahasra Ranjan

Email · Homepage · LinkedIn · Google Scholar

Research interests

Foundation models, interpretability, and robust generalization. Why do large models work, where do they fail, and how do they retain knowledge? I study LLM behavior, knowledge representation, and designing systems that fail gracefully. My work spans NLP (code-switching, multilingual models) and computational biology (protein-molecule interactions).

Education

Indian Institute of Technology Bombay — Mumbai, India
B.Tech in Computer Science and Engineering | GPA: **9.05**/10 | Honors GPA: **9.75**/10 — *Aug 2019 - May 2023*

Honors & awards

- **ACL 2023 Outstanding Paper Award**, first author; awarded to less than 1% of accepted papers — *2023*
- **Research Excellence Award**, best Bachelor's Thesis among graduating CS students, IIT Bombay — *2023*
- **JEE Advanced AIR 127** / 1.2M candidates; KVPY Fellowship; NTSE Scholar — *2017-2019*

Publications

1. **Sahasra Ranjan**, Richeek Das, Shreya Pathak, Preethi Jyothi. *Improving Pretraining Techniques for Code-Switched NLP*. **ACL 2023** Outstanding Paper Award
   Proposed boundary-aware masking strategies (Switch-MLM, Freq-MLM) that improved multilingual model performance on code-switched text by 5-7 F1 points across Hindi-English, Spanish-English, Tamil-English, and Malayalam-English benchmarks.
2. Alexander Kroll, **Sahasra Ranjan**, Martin K. M. Engqvist, Martin J. Lercher. *A general model to predict small molecule substrates of enzymes based on machine and deep learning.* **Nature Communications, 2023** *(160+ citations)*
   Developed ESP, the first general enzyme-substrate prediction model combining protein language models with molecular fingerprints; deployed as public web server (esp.cs.hhu.de) used by researchers in drug discovery and metabolic engineering.
3. Alexander Kroll, **Sahasra Ranjan**, Martin J. Lercher. *ProSmith: Drug-target interaction prediction using a multi-modal transformer network.* **PLoS Computational Biology, 2024**

Research experience

ProSmith: Drug-Target Interaction via Multimodal Transformers — Summer 2023
Research Collaboration | Prof. Martin J. Lercher, Alexander Kroll — *HHU Düsseldorf, Germany*

- Extended enzyme-substrate work to drug-target interaction; designed ProSmith architecture combining bidirectional cross-attention between protein and molecule representations with gated fusion layers.
- Introduced contrastive pre-training on 500k protein-ligand pairs from BindingDB, significantly improving generalization to unseen drug scaffolds, novel binding sites, and out-of-distribution targets.
- Achieved **94.2% accuracy** on held-out test set (30% error reduction over prior state-of-the-art); deployed public tool at deepmolecules.org serving **50k+ predictions** to researchers worldwide.
- Published in **PLoS Computational Biology** (2024); model outperforms existing methods on kinase inhibition prediction and enzyme-substrate benchmarks.

Gradient-based Optimization for Continual Learning — Jan 2023 - May 2023
B.Tech Thesis | Advisors: Prof. Preethi Jyothi, Prof. Manoj Prabhakaran — *IIT Bombay*

- Developed **CommitteeSGD**: maintains a buffer of prior examples and modifies current gradients to penalize updates that conflict with previously learned knowledge, reducing catastrophic forgetting.
- Exploited local elasticity property of neural networks (ICLR'20) to minimize decision boundary perturbation for new examples while preserving classification of old examples.
- Achieved **3.5% accuracy improvement** on CIFAR-100 (VGG16) over standard SGD; periodic committee refresh outperformed static buffer (**66.8%** vs 63.4% test accuracy).

Improving Pretraining for Code-Switched NLP — May 2022 - Dec 2022
B.Tech Thesis | Advisor: Prof. Preethi Jyothi, CSALT Lab — *IIT Bombay*

- Investigated why multilingual models (mBERT, XLM-R) underperform on code-switching; identified that standard MLM ignores language boundaries, causing models to learn spurious monolingual correlations.
- Designed **Switch-MLM**: boundary-aware masking that preferentially masks tokens at language switch-points, forcing models to learn cross-lingual dependencies rather than relying on monolingual context.

- Introduced **Freq-MLM** for low-resource settings without language-ID labels; achieved **5-7 F1 improvement** on NER and sentiment analysis across Hindi-English and Spanish-English benchmarks.
- Published at **ACL 2023** and received **Outstanding Paper Award** (top 1% of submissions); work establishes new best practices for pretraining multilingual models on code-switched corpora.

Deep Learning for Enzyme-Substrate Prediction Summer 2021
Research Internship | Prof. Martin J. Lercher, Alexander Kroll *HHU Düsseldorf, Germany*

- Built one of the first multimodal transformer for enzyme-substrate prediction, jointly encoding protein sequences via ESM-1b embeddings and molecular structures via ECFP fingerprints with learned attention.
- Replaced frozen ESM-1b with end-to-end fine-tuning; improved accuracy from **82% to 91%** by allowing protein representations to adapt dynamically to substrate binding context.
- Curated training data from BRENDA database (70k enzyme-substrate pairs); handled severe class imbalance via strategic negative sampling from biochemically plausible non-substrate molecules.
- Published in **Nature Communications** (160+ citations); model deployed as public web tool at esp.cs.hhu.de, enabling researchers worldwide to predict enzyme-substrate relationships for metabolic engineering and drug discovery applications.

Entrepreneurship

Mathkraft 2024 - 2025
Co-founder, AI Math Tutor for K-8 *San Francisco, CA*

- Co-founded AI tutoring startup to democratize math education; built platform where kids create adventures and characters with math woven into storylines. Piloted with **150+ students** achieving **95%+ retention** over 6 months.
- Managed team of 4 engineers and 6 tutors; conducted 50+ hands-on sessions to understand what kids and parents care about. Discovered younger students need emotional connection before cognitive engagement.
- Human-in-loop model proved unscalable; led to insight that some gaps between models and humans are structural, not accidental, motivating deeper interest in AI alignment and robustness.

Industry experience

Socratic AI Inc Aug 2025 - Present
Founding AI Engineer (2nd engineer) *Sunnyvale, CA*

- Architected Claude Agent SDK integration: designed unified subagent system with task delegation, native streaming, and 70/30 message compaction for long-running conversations. Core infrastructure powering all AI features.
- Built Odoo ERP connector from scratch: schema introspection, field customization agents with infinite loop prevention, CSV validation, and data migration workflows. Reduced 6-week enterprise migrations to hours.
- Designed streaming citation system for Knowledge Q&A: grounding parser, real-time inline references, and artifact resolution. Built full-stack end-to-end (database migrations, backend services, frontend UI).
- Implemented AI-enhanced PDF processing: Claude vision + Gemini for image descriptions, two-pass extraction (instant feedback + async deep analysis), and original-quality image extraction from complex documents.

Rubrik Inc Summer 2022 (Intern), 2023 - 2024 (Full-time)
Cloud Data Software Engineer, Data Life Cycle Team *Bengaluru, India*

- Designed garbage collection for snapshot metadata; built reference tracking, mark-sweep lifecycle, and gRPC APIs for automated orphaned entry detection and removal across distributed storage nodes.
- Extended retention lock and legal hold to cloud-native workloads; implemented archival/replication constraints for immutable backups protecting petabytes of enterprise data.
- Implemented async SLA assignment with event-driven architecture and operation poller for non-blocking parallel workload scheduling across thousands of protected objects.

Teaching

Teaching Assistant, CS725: Foundations of Machine Learning (Prof. Preethi Jyothi) *Fall 2022*
Teaching Assistant, CS753: Automatic Speech Recognition (Prof. Preethi Jyothi) *Spring 2023*

Technical skills

Languages: Python, C++, Go, TypeScript, SQL **ML/DL:** PyTorch, TensorFlow, Hugging Face Transformers, JAX
Infrastructure: Docker, Kubernetes, Git, PostgreSQL, MongoDB, Redis