

Ananjan Nandi

✉ ananjan@stanford.edu

in ananjan-nandi-968386201

🌐 ananjan-nandi-9.github.io

Education

- 2025 – Present • **Ph.D. in Computer Science, Stanford University**
Rotating with Prof. Monica Lam, Stanford NLP Group
- 2023 – 2025 • **M.S. in Computer Science, Stanford University**
Artificial Intelligence Specialization
CGPA: **4.147/4**
- 2019 – 2023 • **B.Tech. in Computer Science and Engineering, Indian Institute of Technology Delhi**
Department Specialization in Data Analytics and Artificial Intelligence
CGPA: **9.877/10** (Institute Rank **4** in cohort of 1000+)

Employment History

- 2024 – Present • **Graduate Research Assistant. Stanford Natural Language Processing (NLP) Group.**
Supervised by Prof. Christopher D. Manning and Prof. Monica Lam.
- 2024 • **ML Researcher Intern. Palantir Technologies.**
Trained a semantic parser for an SQL-like query language, with 20-pt gains over GPT-4.
Built an enterprise Copilot that achieved 25-pt CodeBLEU improvements over GPT-4.
- 2023 • **AI Researcher Intern. KnowDis Data Science.**
Delivered six projects in the fields of multilingual natural language processing, extreme classification and drug discovery.
- 2022 • **Member Technical Intern. D. E. Shaw India Pvt Ltd.**
Achieved up to 4X reduction in response latency for firmwide web services, and sped up the calculation of Value-at-Risk from terabyte-scale profit and loss data by up to 10X.

Selected Research Publications

- 1 **Ananjan Nandi**, Christopher D. Manning, and Shikhar Murty. “Sneaking Syntax into Transformer Language Models with Tree Regularization”. In: *2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics*. 2025.
- 2 Moussa Koulako Bala Doumbouya, **Ananjan Nandi**, Davide Ghilardi, Gabriel Poesia, Anna Goldie, Federico Bianchi, Dan Jurafsky, and Christopher D Manning. “h4rm3l: A Language for Composable Jailbreak Attack Synthesis”. In: *The Thirteenth International Conference on Learning Representations*. 2025.
- 3 **Ananjan Nandi**, Navdeep Kaur, Parag Singla, and Mausam . “DynaSemble: Dynamic Ensembling of Textual and Structure-Based Models for Knowledge Graph Completion”. In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*. 2024.
- 4 Ryan Louie, **Ananjan Nandi**, William Fang, Cheng Chang, Emma Brunskill, and Diyi Yang. “Roleplay-doh: Enabling Domain-Experts to Create LLM-simulated Patients via Eliciting and Adhering to Principles”. In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. 2024.
- 5 **Ananjan Nandi**, Navdeep Kaur, Parag Singla, and Mausam . “Simple Augmentations of Logical Rules for Neuro-Symbolic Knowledge Graph Completion”. In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. 2023.

Projects

- 2025 – Present • **Building a Knowledge Graph for Biomedicine**
PI: Prof. Monica Lam (Stanford Open Virtual Assistant Lab)
Developing automated data curation methods for the creation of a high-quality knowledge graph grounded in PubMed abstracts, with a focus on drug repurposing.
- 2023 – 2024 • **Encouraging Hierarchical Computation in Large Language Models (LLMs)**
PI: Prof. Christopher D. Manning (Stanford NLP Group)
Developed a structured regularization loss to inject syntactic inductive biases into LLMs, improving generalization and out-of-distribution language understanding without changes to the transformer architecture (Published at **NAACL 2025**).

Projects (continued)

- 2024
- **Distributionally Robust Optimization (DRO) for Multilingual Speech Recognition**
PI: Profs. Dan Jurafsky, Karen Livescu, Tatsunori Hashimoto (Stanford NLP Group)
Developed an optimization algorithm to improve performance by up to 10% for worst-performing languages while training ASR models (Under review for **NeurIPS 2025**).
 - **A Language for Composable Jailbreak Attack Synthesis**
PI: Profs. Christopher D. Manning, Dan Jurafsky (Stanford NLP Group)
Proposed a domain-specific language for synthesizing jailbreak attacks at scale, achieving over 90% success rates against LLMs such as Claude-3 (Published at **ICLR 2025**).
- 2023 – 2024
- **Large Language Models in Psychotherapy**
PI: Prof. Diyi Yang (Social and Language Technologies Lab, Stanford)
Built an LLM-based system enabling domain experts to author realistic AI patients to be used in roleplay practice for novice therapists (published at **EMNLP 2024**).
Developed an LLM-based therapist aligned with the Motivational Interviewing framework, whose responses were favored over those of human therapists by expert annotators.
- 2022 – 2023
- **Augmentation and Ensembling Techniques for Knowledge Graph Completion**
PI: Profs. Mausam, Parag Singla (Data Analytics and Intelligence Research Lab, IIT Delhi)
Obtained state-of-the-art results on standard datasets by leveraging a dynamic mixture-of-experts approach to unify structure and text-based KGC methods (published at **ACL 2024**).
Designed simple and performant rule augmentation and pruning techniques for Neuro-Symbolic Knowledge Graph Completion (KGC) (published at **ACL 2023**).
 - **Land Cover Classification from Satellite Data**
PI: Prof. Aaditeshwar Seth (Appropriate Computing Technologies Lab, IIT Delhi)
Developed a pipeline using satellite data from Google Earth Engine for pixel-level land use land cover classification, enhancing existing spatial classifiers (published at **ICTD 2024**).

Proficiencies

- | | |
|------------|--|
| Languages | • Python, C, C++, Java, SML, HTML, JavaScript, CUDA, Bash, \LaTeX |
| AI/ML | • PyTorch, PyTorch-Geometric, PyTorch-Lightning, HuggingFace, LangChain, VLLM, Weights and Biases, Scikit-learn, FAISS, Tslearn, NumPy, Pandas, Deepspeed, Neuron |
| Tools | • Git, Vim, Jupyter, Dask, FastAPI, StreamLit, AsyncIO, Joblib, OpenMP, Google Earth Engine |
| Coursework | • Data Structures and Algorithms, Parallel Programming, Principles of Artificial Intelligence, Machine Learning, Natural Language Processing, Deep Multi-Task and Meta-Learning, Machine Learning with Graphs, Spoken Language Processing, Data Mining (<i>A+ or A in all</i>) |

Academic Achievements

- 2023
- **Outstanding Project Award**, CS 330 (Deep Multi-Task and Meta Learning), Stanford
 - **Graduate Record Examinations: 338/340 (170 - Quantitative, 168 - Verbal)**, ETS
 - **Test of English as a Foreign Language: 119/120**, ETS
- 2022 – 2023
- **Endowment Merit Scholarship**, Indian Institute of Technology Delhi Endowment Fund
- 2019 – 2023
- **Top 7% Merit Prize (Semesters 1, 2, 5, 7 and 8)**, Indian Institute of Technology Delhi
- 2019
- **All India Rank 73 (General Category)**, Joint Entrance Examinations (Advanced)
 - **All India Rank 100 (General Category)**, Joint Entrance Examinations (Mains)
 - **One of 5 selected for the Indian national team**, Asian Physics Olympiad
 - **One of 35 shortlisted for the Indian national team**, International Physics Olympiad

Extracurricular Activities

- 2023 - Present
- **Peer Reviewer.** ACL Rolling Review, NeurIPS, ICLR
- 2023
- **Teaching Assistant. An Introduction to Artificial Intelligence.** NPTEL
- 2021 – 2023
- **Vice Captain. Table Tennis.** Zanskar House, Indian Institute of Technology Delhi
- 2021
- **Academic Mentor. Introduction to Computer Science.** IIT Delhi