Ananjan Nandi

ananjan@stanford.edu ananjan-nandi-968386201 scholar.google.com/user=I3uNMUEAAAAJ

Education

2023 – 2025 • **M.S. in Computer Science, Stanford University**
Artificial Intelligence Specialization (Distinction in Research)
CGPA: **4.124/4**

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 more than 1000)

Employment History

2024 – Present • **Graduate Research Assistant. Stanford Natural Language Processing (NLP) Group.**
Working on projects related to useful structural inductive biases and safety in the context of large language models (supervised by Professor Christopher D. Manning).

2024 • **ML Researcher Intern. Palantir Technologies.**
Developed a model for converting natural language queries into executable queries in an internal query language, delivering over 20-pt performance gains over GPT-4.
Built an enterprise Copilot that achieved 25-pt CodeBLEU improvements over GPT-4.

2023 • **AI Researcher Intern. KnowDis Data Science.**
Executed and delivered six projects in the areas of natural language processing, recommender systems, and molecular AI, all ultimately deployed to production.

2022 • **Member Technical Intern. D. E. Shaw India Pvt Ltd.**
Scaled up firmwide web services serving risk assessment data, achieving up to 4X reduction in response latency under high concurrent request load.
Sped up the calculation of Value-at-Risk from terabyte-scale profit and loss data by up to 10X.

Research Publications

1 **Ananjan Nandi**, Christopher D. Manning, and Shikhar Murty. "Sneaking Syntax into Transformer Language Models with Tree Regularization". In: *Review at ACL Rolling Review* (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: *Review at International Conference of Learning Representations (ICLR)* (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 (ACL 2024)*. Bangkok, Thailand, 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 (EMNLP 2024)*. Miami, USA, 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 (ACL 2023)*. Toronto, Canada, 2023.

Projects

2023 – Present • **Encouraging Hierarchical Structure in Large Language Models (LLMs)**
PI: Professor Christopher D. Manning (Stanford NLP Group)
Developed a structured regularizer usable during both pretraining and finetuning to inject syntactic inductive biases into LLMs, improving generalization and out-of-distribution language understanding without additional parameters, inference complexity or changes to the transformer architecture (Under peer review).

Projects (continued)

2024 • **Evaluating LLM Safety against Composable Jailbreak Attacks**
PI: Professor Christopher D. Manning, Professor Dan Jurafsky (Stanford NLP Group)
Proposed a domain-specific language for synthesizing jailbreak attacks at scale, achieving over 90% success rates against several top LLMs, including Claude-3 (Under peer review).

2023 – 2024 • **Large Language Models in Psychotherapy**
PI: Professor 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 by expert annotators over those of human therapists.

2022 – 2023 • **Augmentation and Ensembling Techniques for Knowledge Graph Completion**
PI: Professor Mausam (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: Professor Aaditeshwar Seth (Appropriate Computing Technologies Lab, IIT Delhi)
Developed a pipeline using temporal satellite data from Google Earth Engine for pixel-level land use land cover classification, enhancing existing spatial classifiers.
In collaboration with 4 NGOs and over 15 experts, validated models with groundtruth data and deployed them in a community-mapping app used by state governments and NGOs to monitor deforestation and cropland usage.

Skills

Languages • Python, C++, Java, SML, HTML, JavaScript, Bash, LaTeX

AI/ML • PyTorch, PyTorch-Geometric, PyTorch-Lightning, HuggingFace, LangChain, VLLM, Weights and Biases, Scikit-learn, FAISS, Tslearn, NumPy, Pandas, Deepspeed

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 *(Highest course grade in each)*

Academic Achievements

2024 • **Outstanding Project Award**, CS 224S (Spoken Language Processing), Stanford

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: 118/120**, ETS

2022 – 2023 • **Endowment Merit Scholarship**, Indian Institute of Technology Delhi Endowment Fund

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