Ananya Sai B

✉ ananya.b.sai@gmail.com   in ananyasaib   🌐 https://ananyasaib.github.io

Research Interests

☐ Natural Language Processing ☐ Machine Learning ☐ Responsible AI ☐ AI for Social Good

◇ Evaluation of Generative AI ◇ Adversarial Attacks ◇ Fairness ◇ Multi agents ◇ Interpretability

Employment History

Oct 2022 – Jan 2023	**Research Intern, Google Brain, UK,** Proposed novel techniques and dimensions for evaluating LLMs for code generation. As a side-project, created and coded a prototype of text-to-video generation application to fuel the imagination and creativity of kids, which won the Special mention award at 'London Hack Week'.
May 2021 – Sep 2021	**Research Intern, Google, India,** Modelled native language influences on acquired languages to address new-age queries from users of diverse backgrounds. This work is live in production at Google and published at ACL-23.
Dec 2018	**Project Officer, IIT Madras, India,** Led a funded project to formulate better dialogue evaluation models and dialogue datasets.
Mar 2018 - Nov 2018	**Research Intern, NVIDIA, India,** Developed techniques to identify vulnerabilities and shortcomings of the existing learnable models used for evaluation of dialogue agents and AI chatbots. (published at AAAI-19)
Aug 2015 – Jun 2016	**Software Engineer, Cisco Systems, India,** Created a data model to enable coordinating large set of routers centrally in the OSPF routing protocol. Improved the network failure-detection ability of RCMD (Route Convergence Monitoring and Diagnostics)

Education

2019 – 2023	**Ph.D., Indian Institute of Technology, Madras** Thesis title: *Towards better evaluation of Natural Language Generation systems*
2016 – 2018	**M.Tech. Computer Science, Indian Institute of Technology, Madras** Thesis title: *Conversation and QA neural network systems to improve Human Machine Interaction.*
2011 – 2015	**B.E. Computer Science, SJCE, Mysore** Thesis title: *Query plan Selector for Optimizing MongoDB.*

Research Publications

Journal Articles

- **Ananya B. Sai**, Akash Kumar Mohankumar, and Mitesh M. Khapra, "A survey of evaluation metrics used for nlg systems," *ACM Computing Surveys*, 2022. 🔗 URL: https://dl.acm.org/doi/full/10.1145/3485766.

- **Ananya B. Sai**, Akash Kumar Mohankumar, and Mitesh M. Khapra, "Improving dialog evaluation with a multi-reference adversarial dataset and large scale pretraining," *TACL*, 2020. 🔗 URL: https://transacl.org/ojs/index.php/tacl/article/view/2389.

Conference Proceedings

- **Ananya B. Sai**, Tanay Dixit, Vignesh Nagarajan, *et al.*, "IndicMT Eval: A Dataset to Meta-Evaluate Machine Translation Metrics for Indian Languages," in *The 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, Toronto, Canada, 2023. 🔗 URL: https://aclanthology.org/2023.acl-long.795.pdf.

- Gupta Abhirut, **Ananya B. Sai**, Richard Sproat, and Aravindan Raghuveer, "Bi-Phone: Modeling Inter Language Phonetic Influences in Text," in *The 61ˢᵗ Annual Meeting of the Association for Computational Linguistics (ACL)*, Toronto, Canada, 2023. 🔗 URL: https://aclanthology.org/2023.acl-long.145.pdf.

- **Ananya B. Sai**, Tanay Dixit, Dev Yashpal Sheth, Sreyas Mohan, and Mitesh M. Khapra, "Perturbation CheckLists for Evaluating NLG Evaluation Metrics," in *The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Punta Cana, Dominican Republic, 2021. 🔗 URL: https://aclanthology.org/2021.emnlp-main.575.pdf.

- **Ananya B. Sai**, Mithun Das Gupta, Mitesh M. Khapra, and Mukundhan Srinivasan, "Re-evaluating ADEM: A Deeper Look at Scoring Dialogue Responses," in *The 33ʳᵈ AAAI Conference on Artificial Intelligence (AAAI)*, Hawaii, USA, 2019. 🔗 URL: https://ojs.aaai.org/index.php/AAAI/article/view/4581.

- Soham Parikh, **Ananya B. Sai**, Preksha Nema, and Mitesh M. Khapra, "ElimiNet: A Model for Eliminating Options for Reading Comprehension with Multiple Choice Questions," in *the 27ᵗʰ International Joint Conference on Artificial Intelligence (IJCAI)*, Stockholm, Sweden, 2018. 🔗 URL: https://www.ijcai.org/proceedings/2018/0594.pdf.

Awards and Achievements

2019 – 23	**Google PhD Fellowship**
2020 – 24	**Prime Minister Fellowship for Doctoral Research**
2023	**Her Research – Our Future award** for a top 5 poster at Springer nature Research Conclave
2021	**NAACL-21 tutorial on Evaluation Metrics used in Natural Language Generation** (co-presented with Dr. Mitesh M. Khapra)
2022	**Bro. C. Selvam Endowment Prize**, as the best performing PhD scholar in the department of Computer Science, IIT Madras
2019	**Star Teaching Assistant** award for Pattern Recognition and Machine Learning course
2018	**Microsoft Research travel grant** and **Kris Gopalakrishnan travel grant** for AAAI-19 and IJCAI-18 conferences
2016	**All India Rank 129 (99.88 percentile)**, in Computer Science, Graduate Aptitude Test in Engineering (GATE)
2006	**Ranked 4th in the state** in class 10ᵗʰ exams

Positions of Responsibility

2018 – 2022	**Teaching Assistant, IIT Madras,** for the courses: Pattern Recognition and Machine Leaning, Deep Learning, Probability and Computing, and Linear Algebra
2018 –	**Conference duties,** Reviewer for ACL 2023, AAAI 2022, ACL 2020 and AAAI 2019 Volunteered to organize IJCAI 2018, ICLR 2020 (virtual), AAAI 2021 (virtual), ACL 2022, ACL 2023
2014 – 2015	**Technical Co-ordinator of Linux Campus Club, SJCE,** Created and hosted coding contests, tech quiz, hackathon to bring the student community with common interests in Coding, Linux and Computer Science together via engaging discussions and collaborative projects.
	Vice-Chairman, Web Designing Board, IEEE-SJCE Organised web designing workshop including conducting sessions on HTML, PHP and evaluating the final contest. Also coded and hosted multiple online treasure hunt events which had to be *hack-proof*

Tutorials and Invited Talks

- NAACL-2021 tutorial on "Evaluation Metrics used in Natural Language Generation" (co-presented with Dr. Mitesh M. Khapra)
- Lecture on "Introduction to Neural Networks, Perceptrons, Multilayer perceptrons, Feedforward neural network" at ACM India Summer Schools 2022
- Lecture on "AI in NLP" to introduce the concepts in a fun and relatable way to high school children
- Talk on "The Paradox of Progress in Automatic Text Generation: Balancing Advancements and Limitations" at Paradox'23, IIT Madras
- Talk titled "A Deeper look at automatic scoring of Dialogue Responses" at the 11th RBCDSAI workshop on Recent progress in Data Science and AI

Skills

Languages	Python, C++, Java, C, R, MATLAB LaTeX
Libraries	Pytorch, Tensorflow, Theano, NLTK
Databases	MySQL, MongoDB

Certification

May 2013	**Artificial Intelligence**, University of California, Berkeley through edX
Oct 2013	**Social Psychology**, Department of Psychology, Wesleyan University through Coursera
Feb 2013	**Advanced Power Searching** with Google

Key Projects

Jun 2017 – Jun 2018

Conversation and QA neural network systems to improve Human Machine Interaction
M. Tech Project, Guide: Dr. Mitesh M. Khapra, (Python, Theano)
– Evaluated the contribution of various attention modules and components in the existing neural network models for QA
– Created simpler versions of the RACE dataset with hints to the answer in the passage or question, to evaluate the components that are supposed to be sensitive to such cues. Proposed and developed a new model, Eliminet, that can eliminate multiple options iteratively.

Jan 2015 – May 2015

Query plan Selector for Optimizing MongoDB
B. Tech Project, Guide: Prashantha Naduthota, (C++, MongoDB)
– Implemented a query plan selector/pruner for MongoDB, to reduce the number of query plans to be executed by
1. directly predicting the best query plan for the given query, whenever possible
2. pruning out bad plans without having to execute them.

References

Available on Request