

# Neeraj Varshney

Ph.D. Candidate (5<sup>th</sup> Year)  
Computer Science (NLP/NLU)  
Arizona State University

Email: [nvarshn2@asu.edu](mailto:nvarshn2@asu.edu)  
Website: [nrjvarshney.github.io](https://nrjvarshney.github.io)  
Semantic Scholar: [Neeraj-Varshney](#)  
LinkedIn: [neerajvarshney97](#)

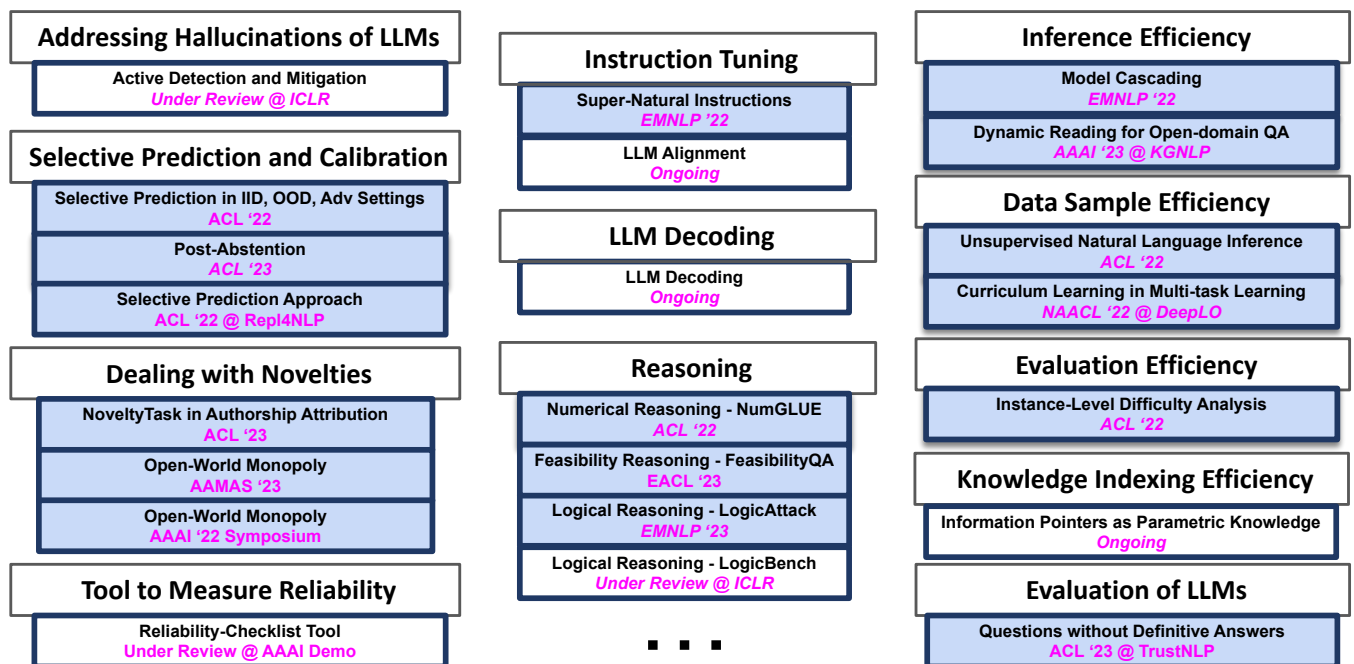
## RESEARCH STATEMENT

I work in Natural Language Processing /Understanding, primarily towards addressing different Efficiency & Reliability aspects of NLP systems in the areas of Open-domain Question Answering, Large Language Models, Information Retrieval, Natural Language Inference, Prompting, and Retrieval Augmented Inference.

On the reliability topic, I focus on Mitigating Hallucinations of LLMs, Selective Prediction & Calibration, NL Generation, Reasoning, Instruction Tuning, and LLM Decoding.

On the efficiency topic, I focus on improving the Inference Efficiency of systems, Question Answering Reader Efficiency, Knowledge Indexing Efficiency, Data Sample Efficiency, and Evaluation Efficiency.

### My Research Work



**Publication Venues:** ACL ('22 & '23), EMNLP ('22 & '23), EACL ('23), NAACL ('22), AAAI ('22 & '23), AAMAS ('23)  
**Thesis Committee:** Dr. Chitta Baral (Chair) Dr. Yezhou Yang Dr. Nakul Gopalan Dr. Pratyay Banerjee

## TECHNICAL SKILLS

**Languages** : Python, Java, C, C++  
**Libraries & Tools** : PyTorch, PyTorch-lightning, Huggingface Transformers, Spacy, OpenAI, Pyserini, NumPy, Matplotlib, Pandas, NLTK, Word2vec, Git, Linux, Amazon Mechanical Turk, PyCharm, Jupyter, Colab, MS Office

## SELECTED PROJECTS

- 1. Detecting and Mitigating Hallucinations of LLMs by Validating Low-Confidence Generation** Preprint, 2023  
*Neeraj Varshney, Wenlin Yao, Hongming Zhang, Jianshu Chen, Dong Yu*
- Addressing the critical problem pertaining to hallucinations of LLMs, we developed an approach that actively detects and mitigates hallucinations during the generation process.
  - In our approach, we first identify the candidates of potential hallucination leveraging the model's logit output values, check their correctness through a validation procedure, mitigate the detected hallucinations via prompting, and then continue generating the subsequent sentences.
  - Showed the effectiveness of our approach in mitigating hallucinations of models such as GPT-3.5 and Vicuna in multiple tasks, such as article generation, multi-hop QA, and false premise QA.

## 2. On Efficiently Indexing External Knowledge for Knowledge Intensive Language Tasks

Under Review, 2023

- Bypassing the requirement of storing vector embeddings of passages of a knowledge corpus and computing a similarity score with query embedding for retrieving relevant knowledge, we developed an approach to index the passages of the corpus in the parameters of a Language Model.
- Trained a generative model to take a query as input and generate identifiers of the passages (from the corpus) that are relevant to the query. Our identifiers also encode the semantic meaning of the passages.

## 3. Post-Abstention: Towards Reliably Re-Attempting the Abstained Instances in QA

ACL, 2023 (Oral)

*Neeraj Varshney, Chitta Baral*

- Developed Post-Abstention methods such as Re-Examining the top-N Predictions (REToP) and an ensembling-based technique that aim at re-attempting to answer the abstained instances of a given selective prediction system with the objective of increasing its ‘coverage’ without significantly sacrificing its ‘accuracy’.
- Showed that state-of-the-art models, even when they are wrong, are often able to rank the ground truth answer as one of their top-N predictions. Building up on this, we developed an auxiliary model that re-examines the top-N predictions of the model to find the correct answer.
- Showed that our approach successfully reduces the risk of the system in both in-domain and out-of-domain settings.

## 4. Dynamic Reading Approach for Efficiently Utilizing External Knowledge in Open-domain QA

AAAI @ KGNLP 2023

*Neeraj Varshney, Man Luo, Chitta Baral*

- Developed an approach that dynamically reads the external knowledge in multiple ‘knowledge iterations’ instead of using a large fixed number of passages for answering open-domain questions.
- Our approach utilizes both the ‘closed-book’ (parametric knowledge) and the ‘open-book’ (external knowledge) inferences in an efficient manner to answer an open-domain question.
- Comparing with the state-of-the-art Fusion-in-Decoder (FiD) reader, our approach matches FiD’s accuracy by utilizing just 18.32% of its reader inference cost (FLOPs) and also outperforms it by achieving up to 55.10% and 77.32% accuracy on NQ Open and TriviaQA respectively.

## 5. Model Cascading: Towards Jointly Improving Inference Efficiency and Accuracy of NLP Systems

EMNLP, 2022

*Neeraj Varshney, Chitta Baral*

- Developed a cascading technique that utilizes a collection of models of varying capacities to accurately yet efficiently output predictions.
- Our methods leverage MaxProb and Distance-to-Uniform values to decide when the prediction with low-cost models is sufficient and when bigger (and relatively higher cost) models are required.

## 6. Investigating Selective Prediction Approaches Across Several Tasks in IID, OOD, and Adv. Settings

ACL, 2022 (F)

*Neeraj Varshney, Swaroop Mishra, Chitta Baral*

- Selective Prediction enables the models to abstain from answering when their prediction is likely to be incorrect; thus improving their reliability. We systematically studied ‘selective prediction’ approaches in a large-scale setup of 17 datasets across NLI, QA, and Duplicate Detection tasks under in-domain, out-of-domain, and adversarial settings.
- Demonstrated that despite leveraging additional resources (such as held-out data or computation), none of the existing approaches consistently and considerably outperforms the simple *MaxProb* baseline. Also analyzed approaches on their task-transfer ability.

## 7. ILDAE: Instance-Level Difficulty Analysis of Evaluation Data

ACL, 2022

*Neeraj Varshney, Swaroop Mishra, Chitta Baral*

- Developed a method to compute instance-level difficulty score for evaluation instances and demonstrated their five novel applications in:
  - Conducting efficient-yet-accurate evaluations with fewer instances saving computational cost and time,
  - Improving the quality of existing evaluation datasets by repairing erroneous and trivial instances,
  - Selecting the best model based on application requirements,
  - Analyzing dataset characteristics for guiding future data creation,
  - Indicating Out-of-Domain performance more reliably.

## 8. Unsupervised Natural Language Inference Using PHL Triplet Generation

ACL, 2022 (F)

*Neeraj Varshney, Pratyay Banerjee, Tejas Gokhale, Chitta Baral*

- Designed three novel unsupervised settings for NLI and proposed a procedural data generation approach that outperforms existing approaches by  $\sim 13\%$  and raises the SOTA unsupervised performance to 66.75%.
- Also developed a general model-in-the-loop adversarial data collection strategy to efficiently collect high-quality non-trivial data instances that help achieve 12.2% higher accuracy with as little as  $\sim 0.1\%$  of the training dataset.

- 9. Instruction Tuning and Benchmarking Generalization on 1,600+ Language Tasks** EMNLP, 2022  
*Yizhong Wang, ..., Neeraj Varshney, ..., Yejin Choi, Hannaneh Hajishirzi, Noah A. Smith, Daniel Khashabi*
- Developed Tk-INSTRUCT, a transformer model trained to follow a variety of in-context instructions (plain language task definitions or k-shot examples).
  - Introduced Super-Natural Instructions, a benchmark of 1,616 diverse NLP tasks and their expert-written instructions.
  - Showed that Tk-INSTRUCT outperforms existing instruction-following models such as InstructGPT by over 9% on our benchmark despite being an order of magnitude smaller.
- 10. NumGLUE: A Suite of Mathematical Reasoning Tasks in NLP** ACL, 2022 (Oral)  
*Swaroop Mishra, Arindam Mitra, Neeraj Varshney, Bhavdeep Sachdeva, Peter Clark, Chitta Baral, Ashwin Kalyan*
- Developed a knowledge-retrieval based multi-task learning method that outperforms existing models.
  - Built a multi-task benchmark that evaluates NLP systems on eight different numerical understanding tasks and evaluated the efficacy of neural models including large language models.
- 11. On Developing Adversarial Attacks for Evaluating Logical Consistency of LLMs** EMNLP, 2023  
*Mutsumi Nakamura, Santosh Mashetty, Mihir Parmar, Neeraj Varshney, Chitta Baral*
- 12. Towards Improving Selective Prediction Ability of NLP Systems** ACL @ RepL4NLP, 2022  
*Neeraj Varshney, Swaroop Mishra, Chitta Baral*
- To improve the selective prediction performance (and hence the reliability) of a system, we developed a method that calibrates the model outputs using prediction confidence and difficulty level of the instances.
  - Instantiated the proposed method in NLI and Duplicate Detection tasks and showed that it outperforms existing approaches and achieves up to 15% improvement over the MaxProb baseline.
- 13. A Unified Evaluation Framework for Novelty Detection and Accommodation in NLP** ACL, 2023 (F)  
*Neeraj Varshney, Himanshu Gupta, Eric Robertson, Bing Liu, Chitta Baral*
- 14. On Dealing with Questions that Don't have Definitive Answers** ACL @ TrustNLP, 2023  
*Neeraj Varshney\*, Ayushi Agarwal\*, Nisarg Patel\*, Mihir Parmar, ..., and Chitta Baral*
- 15. On Evaluating NLP Models' Understanding of Feasibility** EACL, 2023  
*Himanshu Gupta, Neeraj Varshney, Swaroop Mishra, kuntal Pal, S. Sawant, K. Scaria, S. Goyal, Chitta Baral*
- Developed FeasibilityQA, a question-answering dataset involving binary classification and multi-choice multi-correct questions that test understanding of feasibility.
  - Demonstrated that even state-of-the-art models such as GPT-3 struggle to answer the feasibility questions correctly.
- 16. Designing a Learning Curriculum for Developing a Multitask Model** NAACL @ DeepLo, 2022  
*Neeraj Varshney, Swaroop Mishra, Chitta Baral*
- Developed dataset and instance-level techniques to arrange training instances into a learning curriculum based on the model's own interpretation of difficulty.
  - Achieved 4% accuracy improvement over other methods on experiments conducted for 12 datasets covering a variety of language understanding tasks.
- 17. Methods and Mechanisms for Interactive Novelty Handling in Adversarial Environments** AAMAS 2023 (E)  
*Tung Thai, M. Shen, ..., Neeraj Varshney, Chitta Baral, Subbarao Kambhampati, Jivko Sinapov, Matthias Scheutz*

## EXPERIENCE

- |  |   |
|--|---|
| <b>Tencent AI</b><br>NLP Research Intern   | May 2023 – Aug 2023<br>Bellevue, Washington |
| <ul style="list-style-type: none"> <li>Developed an approach for Detecting and Mitigating Hallucinations of Large Language Models.</li> </ul>                          |   |
| <b>Amazon Science</b><br>Applied Scientist Intern, Alexa AI  | May 2022 – Aug 2022<br>Seattle, Washington  |
| <ul style="list-style-type: none"> <li>Web Question-Answering system using Information Retrieval.</li> </ul>   |   |
| <b>Microsoft</b><br>Software Developer   | July 2018 – Aug 2019<br>Bangalore, India    |
| <ul style="list-style-type: none"> <li>An ML driven chat recommendation system aimed at augmenting user engagement with Microsoft 'Teams'.</li> </ul>                  |   |
| <b>Samsung R&amp;D Institute</b><br>Research Intern  | Summer 2017<br>Bangalore, India             |
| <ul style="list-style-type: none"> <li>Developed a 'context prediction' application leveraging event features such as app usage, location, and sensor data.</li> </ul> |   |

## EDUCATION

### Arizona State University

Ph.D. in Computer Science

Tempe, AZ, USA

Fall 2019 – Spring 2024 (Expected)

- **Advisor:** Dr. Chitta Baral
- **CGPA:** 4/4
- **Awards:** SCAI doctoral fellowship, GPSA awards (3 times), SCAI conference award (2 times), Graduate College awards (5 times), AAAI student scholarship, ACL registration award.
- **Internships:** Amazon Science (Summer 2022), Tencent AI (Summer 2023)

### BITS Pilani, Pilani Campus, India

B.E (Hons) Computer Science

Pilani, India

2014-2018

- **CGPA:** 9.11/10 (with Distinction)
- **Experience:** 'Web Intelligence & Social Computing' research lab under Prof. Poonam Goyal, CEERI research lab under Dr. J.L. Raheja.
- **Internships:** Microsoft, Samsung R&D Institute, Valuefirst Digital Media.

## HONORS AND AWARDS

- **Outstanding Reviewer** for EACL'23 (Question Answering track).
- SCAI Doctoral Fellowship, ASU, 2023.
- ASU Jumpstart Research Grant, 2023.
- Selected for AAAI Student Scholarship, 2023.
- Graduate College Travel Award, ASU for AAAI 2023, ACL 2022, NAACL 2022, EMNLP 2022, and ACL 2023.
- GPSA Award, ASU for EMNLP 2022 and ACL 2022.
- SCAI Conference Award, ASU for EMNLP and ACL '22.
- Registration award from Repl4NLP for ACL, 2022.
- GPSA Internship Travel Award, ASU 2023.

## SERVICE

- Reviewer for **ACL'23, EMNLP'23, EACL'23** (Outstanding Reviewer), **CVPR workshop** (Open-Domain Retrieval Under a Multi-Modal Setting) 2022 & 2023.
- Reviewer for GPSA Research Grants, ASU.
- Mentored B.S and M.S students for course projects and co-authored multiple research papers with them.
- Served as Maths teacher for underprivileged kids through National Service Scheme (NSS), India.
- Participated in blood donation camps and health awareness drives.

## COURSES

Natural Language Processing  
Knowledge Representation

Statistical Machine Learning  
Data Mining

Artificial Intelligence  
Social Media Mining

NLP Methods in BioMedical  
Mobile Computing

## COLLABORATORS

- **Swaroop Mishra** (Research Scientist at Google Brain)
- **Tejas Gokhale** (Assistant Professor at University of Maryland, Baltimore County)
- **Arindam Mitra** (Data and Applied Scientist at Microsoft Research)
- **Bing Liu** (Professor at University of Illinois at Chicago)
- **Daniel Khashabi** (Allen AI, Assistant Professor at Johns Hopkins University)
- **Pratyay Banerjee** (Applied Scientist at Alexa AI, Amazon)
- **Kuntal Pal** (Applied AI ML Senior Associate at JPMorgan Chase & Co.)

- **Dong Yu** (Distinguished Scientist at Tencent AI)
- **Jianshu Chen** (Principal Researcher at Tencent AI)
- **Hongming Zhang** (Senior Research Scientist at Tencent AI)
- **Wenlin Yao** (Senior Research Scientist at Tencent AI)
- **Ashwin Kalyan** (Allen AI)
- **Yizhong Wang** (Allen AI, University of Washington)
- **Rik Koncel-Kedziorski** (Alexa AI)
- **Eric Robertson** (PAR Government)
- **Man Luo** (ASU)
- **Mihir Parmar** (ASU)