Neeraj Varshney

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

Email: nvarshn2@asu.edu
Website: nrjvarshney.github.io
Semantic Scholar: Neeraj-Varshney
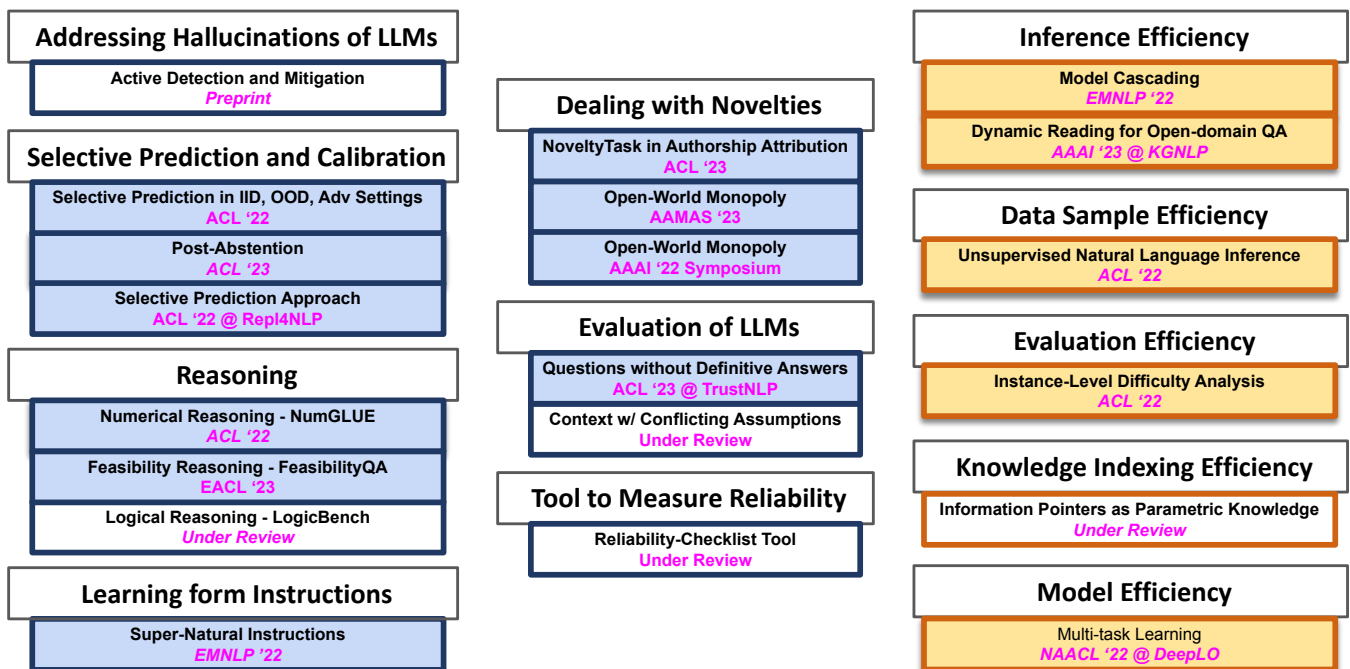LinkedIn: neerajvarshney97

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, and Retrieval Augmented Inference.

On the reliability topic, I focus on Mitigating Hallucinations of LLMs, Selective Prediction & Calibration, Reasoning, and Learning from Instructions.

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

Addressing Hallucinations of LLMs

Active Detection and Mitigation
Preprint

Selective Prediction and Calibration

Selective Prediction in IID, OOD, Adv Settings
ACL '22
Post-Abstention
ACL '23
Selective Prediction Approach
ACL '22 @ Repl4NLP

Reasoning

Numerical Reasoning - NumGLUE
ACL '22
Feasibility Reasoning - FeasibilityQA
EACL '23
Logical Reasoning - LogicBench
Under Review

Learning form Instructions

Super-Natural Instructions
EMNLP '22

Dealing with Novelties

NoveltyTask in Authorship Attribution
ACL '23
Open-World Monopoly
AAMAS '23
Open-World Monopoly
AAAI '22 Symposium

Evaluation of LLMs

Questions without Definitive Answers
ACL '23 @ TrustNLP
Context w/ Conflicting Assumptions
Under Review

Tool to Measure Reliability

Reliability-Checklist Tool
Under Review

Inference Efficiency

Model Cascading
EMNLP '22
Dynamic Reading for Open-domain QA
AAAI '23 @ KGNLP

Data Sample Efficiency

Unsupervised Natural Language Inference
ACL '22

Evaluation Efficiency

Instance-Level Difficulty Analysis
ACL '22

Knowledge Indexing Efficiency

Information Pointers as Parametric Knowledge
Under Review

Model Efficiency

Multi-task Learning
NAACL '22 @ DeepLO

Publication Venues: ACL EMNLP EACL NAACL AAAI AAMAS

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 propose 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, and then continue generating the subsequent sentences.

- We show the efficacy of our approach in multiple tasks, such as article generation task, multihop QA, and false premise QA.

2. On Efficiently Indexing External Knowledge for Open-QA   Under Review, 2023

- We explore how well we can index an external knowledge corpus in the parameters of a Language Model. This bypasses the need to store vector embeddings for the corpus and computing similarity scores with the query embedding at run time.

3. Post-Abstention: Towards Reliably Re-Attempting the Abstained Instances in QA   ACL, 2023 (Oral)

Neeraj Varshney, Chitta Baral

- A Post-Abstention method aims at re-attempting the abstained instances of a given selective prediction system with the objective of increasing its 'coverage' without significantly sacrificing its 'accuracy'.
- We propose several methods such as Re-Examining the top-N Predictions (REToP) and ensembling-based methods.

4. Can Open-Domain QA Reader Utilize External Knowledge Efficiently like Humans?   AAAI @ KnowledgeNLP 2023

Neeraj Varshney, Man Luo, Chitta Baral

- Proposed 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 outperformes 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

- Proposed a cascading technique that utilizes a collection of models of varying capacities to accurately yet efficiently output predictions.

6. Investigating Selective Prediction Approaches Across Several Tasks in IID, OOD, and Adv. Settings   ACL, 2022

Neeraj Varshney, Swaroop Mishra, Chitta Baral

- 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 evaluated approaches on their task-transfer ability.

7. ILDAE: Instance-Level Difficulty Analysis of Evaluation Data   ACL, 2022

Neeraj Varshney, Swaroop Mishra, Chitta Baral

- Computed instance-level difficulty scores for evaluation instances and demonstrated their five novel applications:
  - 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

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 proposed a 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. NumGLUE: A Suite of Mathematical Reasoning Tasks   ACL, 2022 (Oral)

Swaroop Mishra, Arindam Mitra, Neeraj Varshney, Bhavdeep Sachdeva, Peter Clark, Chitta Baral, Ashwin Kalyan

- Proposed a multi-task benchmark that evaluates AI systems on eight different numerical understanding tasks and evaluated neural models including large language models.
- Proposed a knowledge-retrieval based multi-task learning method that outperforms existing models.

10. Towards Improving Selective Prediction Ability of NLP Systems   ACL @ RepL4NLP, 2022

Neeraj Varshney, Swaroop Mishra, Chitta Baral

- Proposed a method that improves over MaxProb by calibrating 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 MaxProb.

11. Benchmarking Generalization via In-Context Instructions on 1,600+ Language Tasks EMNLP, 2022

Yizhong Wang, …, <u>Neeraj Varshney</u>, …, Yejin Choi, Hannaneh Hajishirzi, Noah A. Smith, Daniel Khashabi

- Introduced Super-NaturalInstructions, a benchmark of 1,616 diverse NLP tasks and their expert-written instructions.
- Built Tk-Instruct, a transformer model trained to follow a variety of in-context instructions (plain language task definitions or k-shot examples).

12. A Unified Evaluation Framework for Novelty Detection and Accommodation in NLP ACL, 2023

<u>Neeraj Varshney</u>, Himanshu Gupta, Eric Robertson, Bing Liu, Chitta Baral

13. On Dealing with Questions that Don't have Definitive Answers ACL @ TrustNLP, 2023

<u>Neeraj Varshney</u>, Ayushi Agarwal*, Nisarg Patel*, Mihir Parmar, …, and Chitta Baral*

14. On Evaluating NLP Models' Understanding of Feasibility EACL, 2023

Himanshu Gupta, <u>Neeraj Varshney</u>, Swaroop Mishra, kuntal Pal, S. Sawant, K. Scaria, S. Goyal, Chitta Baral

- Introduced 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.

15. Let the Model Decide its Curriculum for Multitask Learning NAACL @ DeepLo, 2022

<u>Neeraj Varshney</u>, Swaroop Mishra, Chitta Baral

- Proposed 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.

16. An Architecture for Novelty Handling in a Multi-Agent Stochastic Environment AAAI Sym. 2022

Tung Thai, M. Shen, <u>Neeraj Varshney</u>, S. Gopalakrishnan, U. Soni, Matthias Scheutz, Chitta Baral, Jivko Sinapov

- Introduced an architecture that allows agents to detect novelties, characterize those novelties, and build an appropriate adaptive model to accommodate them.

EXPERIENCE

Tencent AI May 2023 – Aug 2023
NLP Research Intern *Bellevue, Washington*

- Detecting and Mitigating Hallucinations of Large Language Models.

Amazon Science May 2022 – Aug 2022
Applied Scientist Intern *Seattle, Washington*

- Web Question-Answering using Information Retrieval for Alexa AI.

Microsoft July 2018 – Aug 2019
Software Developer *Bangalore, India*

- Worked towards developing a machine learning driven chat recommendation system aimed at augmenting user engagement with the product Microsoft 'Teams'.

Samsung R&D Institute Summer 2017
Research Intern *Bangalore, India*

- Developed a 'context prediction' application incorporating features based on device events (e.g. app usage, location) and sensor data (proximity sensor).

EDUCATION

Arizona State University   Tempe, AZ, USA
Ph.D. in Computer Science   *Aug 2019 – Spring 2024 (Expected)*

- **Advisor**: Dr. Chitta Baral
- **CPGA**: 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   Pilani, India
B.E (Hons) Computer Science   *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, School of Computing and AI at ASU, 2023.
- 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 2022 and ACL 2022.
- Registration award from Repl4NLP for ACL, 2022.
- GPSA Internship Travel Award, ASU 2023.

SERVICE

- Reviewer for **EACL'23** (Question Answering track) - Outstanding Reviewer.
- Reviewer for **ACL'23**, **EMNLP'23**, **CVPR workshop** (Open-Domain Retrieval Under a Multi-Modal Setting) '22 & '23.
- 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.

COURSES

Natural Language Processing	Statistical Machine Learning	Artificial Intelligence	NLP Methods in BioMedical
Knowledge Representation	Data Mining	Social Media Mining	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)