Neeraj Varshney

Ph.D. Student (4th 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, primarily in the areas of Open-domain QA, Natural Language Inference, Retrieval augmented inference, addressing different Efficiency & Reliability aspects of NLP systems, and Natural Language Understanding in general.

On the efficiency topic, I have worked on improving inference efficiency, training sample efficiency, QA reader efficiency, knowledge indexing efficiency, and evaluation efficiency. Furthermore, on the reliability topic, I have worked on selective prediction.

Publication Venues: ACL   EMNLP   EACL   NAACL   AAAI   AAMAS.

Research Interests: Efficiency and Reliability aspects in NLP
Open-domain Question Answering
Robustness
Retrieval-Reader Systems
Multi-task Learning
Learning From Natural Language Instructions
Selective Prediction
Generalization
Learning From Less Supervision

TECHNICAL SKILLS

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

SELECTED PROJECTS

On Efficiently Indexing External Knowledge for Open-QA   Under Review @ ACL 2023

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

- Investigated an approach that utilizes both the 'closed-book' (parametric knowledge) and the 'open-book' (external knowledge) inferences in an efficient manner.
- Instead of using a large fixed number of passages for open-book inference, the system dynamically reads the external knowledge in multiple 'knowledge iterations'.
- Comparing with the Fusion-in-Decoder (FiD) reader, our approach matches FiD's accuracy by utilizing just 18.32% of its reader inference cost (FLOPs) and also outperformed it by achieving up to 55.10% and 77.32% accuracy on NQ Open and TriviaQA respectively.

Unsupervised Natural Language Inference Using PHL Triplet Generation   ACL, 2022

- 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 strategy to efficiently collect high-quality task-specific data instances that helps achieve 12.2% higher accuracy with as little as $\sim0.1\%$ of the training dataset.

Investigating Selective Prediction Approaches Across Several Tasks in IID, OOD, and Adversarial Settings ACL, 2022

- Systematically studied 'selective prediction' in a large-scale setup of 17 datasets across NLI, QA, and Duplicate Detection tasks under in-domain, out-of-domain, and adversarial settings.
- Showed that despite leveraging additional resources (held-out data/computation), none of the existing approaches consistently and considerably outperforms *MaxProb*.

- Also evaluated the existing approaches on their task-transfer ability.

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

- Conducted instance-level difficulty analysis in a large-scale setup of 23 datasets with 27 models and demonstrated its five novel applications such as:
  - Conducting efficient-yet-accurate evaluations with fewer instances saving computational cost and time,
  - Improving 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,
  - Estimating Out-of-Domain performance reliably.

NumGLUE: A Suite of Mathematical Reasoning Tasks ACL, 2022

- 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 MTL method that outperforms existing models.

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

- Proposed a method that improves over MaxProb by calibrating the model outputs using prediction confidence and difficulty level of the instances.
- Instantiated our method in NLI and Duplicate Detection tasks and showed that it outperforms existing approaches and achieves up to 15% improvement over MaxProb.

Model Cascading: Towards Jointly Improving Inference Efficiency and Accuracy of NLP Systems EMNLP, 2022

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

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

- 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).

On Evaluating NLP Models' Understanding of Feasibility EACL, 2023

- Introduced FeasibilityQA, a question-answering dataset involving binary classification and multi-choice multi-correct questions that test understanding of feasibility.
- Showed that even state-of-the-art models such as GPT-3 struggle to answer the feasibility questions correctly.

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

- 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 NLU tasks.

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

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

EXPERIENCE

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

- Web Question-Answering System 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 (2 times), SCAI conference award (2 times), Graduate College awards (4 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

- SCAI Doctoral Fellowship, School of Computing and AI, ASU.
- AAAI Student Scholarship 2023
- Graduate College Travel Award, ASU for AAAI 2023, EMNLP 2022, NAACL 2022, and ACL 2022.
- GPSA Travel 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.

SERVICE

- Reviewer for EACL'23 (Question Answering track).
- Reviewer for CVPR'22 workshop (Open-Domain Retrieval Under a Multi-Modal Setting).
- Mentored B.S and M.S students for class projects and co-authored multiple research papers.
- Served as Maths and Science teacher for underprivileged kids in National Service Scheme (NSS).
- Member of blood donation camps and health awareness drives.

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 (ASU)

- Pratyay Banerjee (Alexa AI, Amazon)

- Tejas Gokhale (ASU)

- Man Luo (ASU)

- Arindam Mitra (Microsoft Research)

- Daniel Khashabi (Allen AI, Assistant Professor@Johns Hopkins University)

- Ashwin Kalyan (Allen AI)

- Yizhong Wang (Allen AI, University of Washington)

- Rik Koncel-Kedziorski (Alexa AI)

OTHERS

- Published 20+ ML/NLP-related articles on medium with 80,000+ views.

- Worked with Dr. Ayush Choure and Dr. Prateek Jain (MSR, India)

- Campus Coordinator of Computer Science Association at BITS Pilani and organized numerous on-campus events, such as Alumni Research Talks and Technical Fest competitions.