Neeraj Varshney

Ph.D. Student (Third Year)
Computer Science (NLP)
Arizona State University, CGPA: 4.1/4

Email: **nvarshn2@asu.edu**
Website: **nrjvarshney.github.io**
Mob: **+1(480)561-9416**
Tempe, Arizona, 85281

Profiles

Website: **nrjvarshney.github.io**
LinkedIn: **neerajvarshney97**
Scholar: **Ju9nR0IAAAAJ&hl**
Medium: **@nvarshney97**
Github: **nrjvarshney**

Education

Ph.D. Computer Science (NLP)
Arizona State University, USA
2019-2024 EXPECTED
CGPA: 4/4

B.E.(Hons.) Computer Science
Birla Institute of Technology
and Science (BITS Pilani, India)
2014-2018
CGPA: 9.11/10

Coursework

Natural Language Processing
Statistical Machine Learning
Artificial Intelligence
NLP Methods in BioMedical
Knowledge Representation &
Reasoning, Data Mining

Technical Skills

ML, DL, PyTorch, Transformers,
Pytorch-lightning, Python,
Jupyter, Pandas, Git, GitHub,
Google Colab, Spacy, Linux,
NumPy, Matplotlib, Huggingface,
NLTK, word2vec

OTHERS

● Worked with Dr. Ayush Choure
(MSR) in a project lead by Dr.
Prateek Jain (MSR).
● Published 15+ ML/NLP articles
on medium with 35000+ views.
● Organized 6th edition of Alumni
Research Talks being the Campus
Coordinator of Computer
Science Association at BITS.
● Worked at "Web Intelligence &
Social Computing" lab under
Prof. Poonam Goyal at BITS.
● Won 2nd prize in Technical fest
prototype project at BITS Pilani.

Publications

Unsupervised NLI 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 ~13% and raises the SOTA unsupervised SNLI performance to 66.75% without using any human-authored examples.
● Also proposed a strategy to efficiently crowd-source high-quality task-specific data that helps achieve 12.2% higher accuracy on SNLI with as little as ~0.1% of the training dataset.

Investigating Selective Prediction in IID, OOD, and ADV settings *ACL, 2022*
● Systematically studied 'selective prediction' in a large-scale setup of 17 datasets across NLI, QA, 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:
● Efficient Evaluations: Proposed an instance selection technique that achieves 0.72 Kendall correlation with full dataset evaluation using just 20% evaluation instances.
● OOD Correlation: Proposed a method to compute weighted accuracy leveraging our difficulty scores and showed that it leads to 5.2% higher correlation with OOD performance.

It's Better to Say "I Can't Answer" than to Answer Incorrectly *ARR, 2022*
● Demonstrated that prominently used selective prediction technique 'MaxProb' fails to perform well in out-of-domain setting while achieving good performance in in-domain setting.
● 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.

Let the Model Decide its Curriculum for Multitask Learning *ARR, JAN, 2022*
● Proposed two classes of techniques (dataset and instance-level) to arrange training instances into a learning curriculum based on model's own interpretation of difficulty.
● Achieved 4% accuracy improvement over other methods on experiments conducted for 12 datasets covering varied sentence pair tasks such as NLI and duplicate detection.

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 showed that it is far from being solved with neural models including large language models performing significantly worse than humans (lower by 46.4%).
● Proposed a knowledge-retrieval based MTL method that outperforms existing models.

Interviewer-Candidate Role Play: Towards Real-World NLP Systems *PREPRINT, 2021*
● Designed a multi-stage selective prediction task in which a system is assisted with instance-related information such as knowledge when it is not sufficiently confident in its prediction.
● Achieved improvements of up to 72.02% in Out-of-Domain generalization.

Ongoing Projects

● Study Post-Abstention techniques such as 'test-time adaptation' and 'question rewriting' that help the system answer questions on which it is not sufficiently confident.
● Learning from instructions: Investigate the impact of conditioning large language models like GPT-3 on task instructions and a few examples in natural language format.

Experience

Microsoft *JAN 2018 - JULY 2019* **Software Engineer**
● Contributed towards development of a Machine Learning driven chat recommendation system aimed at augmenting user engagement with Microsoft's product 'Teams'.
● Collaborated with MSR researchers for a feature titled 'Intelligent Feeds' that finds relevant messages for users based on their prior activities and message text features.