

Vivek Gupta

Machine Learning & Natural Language Processing

CONTACT INFORMATION	The Preserve Apartment 531 South 900 East, Apt. A8, Salt Lake City, UT 84102 United States of America	vgupta123.github.io keviv9@gmail.com +1-801-558-7272 Google Scholar
RESEARCH INTERESTS	Semi-structured Data, Structured Prediction, Natural Language Inference, Representation Learning, Extreme Multi-Label Learning, Resource Constrained Machine Learning, Fairness and Ethics in AI	
CURRENT POSITION	School of Computing, University of Utah <i>Graduate Research Assistant</i>	Aug 2018 - Present Advisor: Prof. Vivek Srikumar
EDUCATION	School of Computing, University of Utah PhD, Computer Science, School of Computing Advisor: Prof. Vivek Srikumar (CGPA: 3.9/4) Indian Institute of Technology, Kanpur M. Tech, Computer Science and Engineering Advisor: Prof. Harish Karnick (CGPA: 9.3/10) Indian Institute of Technology, Kanpur B. Tech, Computer Science and Engineering (CGPA: 7.5/10)	Aug 2018 - Present July 2015 - May 2016 July 2011 - July 2015
PREVIOUS EMPLOYMENT	Microsoft Research Lab, India <i>Research Fellow, Machine Learning and Natural Language Application</i> Dr. Praneeth Netrapalli & Dr. Prateek Jain	Oct 2016 - Aug 2018 Dr. Nagarajan Natarajan,
SCHOLASTIC ACHIEVEMENTS	Best paper award at DeeLIO workshop at ACL 2022 Outstanding paper award at NLP4ConvAI workshop at ACL 2022 One of the recipients of the Bloomberg Data Science Fellowship 2021-2022. Renewed till 2023 Selected to attend Virtual-HLF 2020, MLSS-18 Madrid, MLSS-19 London, MLSS-20 Tubingen Selected among the top 7 finalist in Ericson's Innovation Awards, India in 2016 Selected in top 100 all over India in Telnor, Internet for all challenge by Unisys in 2015 Best poster presentation for Internship works at Flipkart Internet Private Ltd. in 2015 2nd Best project for project on <i>QA System</i> in NLP (CS 5340/6340) in 2018 Best project award for project <i>Distributed Classroom</i> in Topic in Distributed System in 2015 Secured All India Rank 183 in IIT-JEE 2011 amongst more than 500,000 candidates Qualified for Kishore Vaigyanik Protsahan Yojana (KVPY) scholarship, 2011 Selected among top 1% of students, Indian National Physics Olympiads, 2011 Selected among top 1% of students, Indian National Chemistry Olympiads, 2011 Selected among top 1% of students in State, Indian National Physics Olympiads in 2010	
CONFERENCE AND JOURNAL PUBLICATIONS	Gupta, V. , Bhat, R., Ghosal, A., Srivastava, M., Singh, M., Srikumar V., <i>Is My Model Using The Right Evidence? Systematic Probes for Examining Evidence-Based Tabular Reasoning</i> , TACL 2022 [Paper] Gupta, V. , Zhang, S., Vempala, A., He, Y., Choji, T., Srikumar V., <i>Right for the Right Reason: Evidence Extraction for Trustworthy Tabular Reasoning</i> , ACL 2022 [Paper] Agarwal, C.*, Gupta, V.* , Kunchukuttan, A., Shrivastava, M., <i>Bilingual Tabular Inference: A Case Study on Indic Languages</i> , NAACL 2022 [Paper]	

- Aggarwal, D.*, **Gupta, V.***, Kunchukuttan, A., *IndicXNLI: Evaluating Multilingual Inference for Indian Languages*, [EMNLP 2022](#), Short version [MIA-2022](#) [\[Preprint\]](#) [\[Paper\]](#)
- Jena, A.*, **Gupta, V.***, Shrivastava, M., Eisenschlos, J., *Leveraging Data Recasting to Enhance Tabular Reasoning*, [EMNLP 2022](#) in Findings, Short version at [SUKI-2022](#) [\[Paper\]](#)
- Kumar, D.*, **Gupta, V.***, Sharma, S., Zhang, S., *Realistic Data Augmentation Framework for Enhancing Tabular Reasoning*, [EMNLP 2022](#) in Findings, Short version at [SUKI-2022](#) [\[Paper\]](#)
- Shankarampeta, A.*, **Gupta, V.***, Zhang, S., *Enhancing Tabular Reasoning with Pattern Exploiting Training*, [AAACL 2022](#), Short version at [SUKI-2022](#) [\[Paper\]](#)
- Jain, N.*, **Gupta, V.***, Rai, A., Kumar, G., *TabPert: An Effective Platform for Tabular Perturbation* [EMNLP 2021](#), Demo track [\[Paper\]](#) [\[Project\]](#)
- Neeraja, J.*, **Gupta, V.***, Srikumar, V. *Incorporating External Knowledge to Enhance Tabular Reasoning*, [NAACL 2021](#) [\[Paper\]](#) [\[Project\]](#)
- Gupta, V.**, Mehta, M., Nokhiz, P., Srikumar, V. *InfoTabS: Inference on Tables as Semi-structured Data*, [ACL 2020](#) [\[Paper\]](#) [\[Project\]](#)
- Gupta, V.**, Saw A., Nokhiz, P., Gupta, H., Talukdar, P., *Improving Document Classification using Multi-Sense Embeddings*, [ECAI 2020](#) , Short version at [NAACL-SRW 2019](#) [\[Paper\]](#) [\[Blog\]](#)
- Gupta, V.**, Kumar, A., Nokhiz, P., Netrapalli, P., Rai, P., Talulkdar, *P-SIF: Document Embeddings using Partition Averaging*, [AAAI 2020](#) [\[Paper\]](#) [\[Appendix\]](#) [\[PPT\]](#) [\[Poster\]](#) [\[Blog\]](#)
- Uppal S., **Gupta, V.**, Swaminathan A., Zhang H., Mahata D., Gosangi R., Shah. R., Stent A., *Two-Stage Classification using Recasted Data for Low Resource Settings*, [AAACL-IJCNLP 2020](#) [\[Paper\]](#)
- Raunak, V., Dalmia, S., **Gupta, V.**, Metze, F., *On Long-Tailed Phenomena in Neural Machine Translation*, [EMNLP 2020](#) (Findings) & presented at [SPNLP 2020](#) [\[Paper\]](#)
- Li, T., **Gupta, V.**, Mehta, M., Srikumar, V., *A logic-Driven Framework for Consistency of Neural Models*, [EMNLP-IJCNLP 2019](#) & [StarAI 2020](#) [\[Paper\]](#) [\[Poster\]](#)
- Gupta, V.**, Wadbude, R., Natarajan, N., Karnick, H., Jain, P., Rai, P., *Distributional Semantics meet Multi-Label Learning*, [AAAI 2019](#) [\[Paper\]](#) [\[Slides\]](#) [\[Poster\]](#)
- Mekala, D.*, **Gupta, V.***, Paranjape, B., Karnick, H. *Sparse Composite Document Vectors using soft clustering over distributional representations*, [EMNLP 2017](#) [\[Paper\]](#) [\[Slides\]](#)
- Gupta, V.**, Karnick, H., Bansal, A., Jhala, P. *Product Classification in E-Commerce using Distributional Semantics*, [COLING 2016](#) (Master Thesis Work) [\[Paper\]](#) [\[Poster\]](#)

*represent equal contribution

WORKSHOPS PUBLICATIONS

- Gupta, V.**, Shrivastava, A., Sagar, A., Aghajanyan, A., Savenkov. D., *RetroNLU: Retrieval Augmented Task Oriented Semantic Parsing*, [Spa-NLP-2022](#) and [NLP4ConvAI-2022](#), (**Outstanding Paper**) [\[Paper\]](#)
- Varun, Y.*, **Gupta, V.***, Sharma, A., *Trans-KBLSTM: An External Knowledge Enhanced Transformer BiLSTM model for Tabular Reasoning*, [DeeLIO 2022](#) (**Best Paper**) [\[Paper\]](#)
- Minhas, B.*, Shankhdhar, A.*, **Gupta, V.***, Aggarwal, D., Zhang, S., *XInfoTabS: Evaluating Multilingual Tabular Natural Language Inference*, [Fever 2022](#) and [MML 2022](#) [\[Paper\]](#)
- Gupta, A., **Gupta, V.**, *Unsupervised Contextualized Document Representation*, [SustainNLP 2021](#) [\[Paper\]](#)
- Gupta, V.**, Bharti P., Nokhiz, P., Karnick, H., *SumPubMed: Summarization Dataset of PubMed Scientific Articles*, [ACL-IJCNLP SRW 2021](#) [\[Preprint\]](#) [\[Dataset\]](#)
- Yadav, P., Yadav, P., Nokhiz, P., **Gupta, V.**, *Unbiasing Review Ratings with Tendency based Collaborative Filtering*, [AAACL-IJCNLP SRW 2020](#) [\[Paper\]](#)
- Raunak, V., **Gupta, V.**, Metze, F., *Effective Dimensionality Reduction for Word Embeddings*, [RepL4NLP 2019](#) [\[Paper\]](#) [\[Poster\]](#)
- Raunak, V., Kumar, V., **Gupta, V.**, Metze, F., *On Dimensional Linguistic Properties of the Word*

	<i>Embedding Space</i> , ACL-SRW 2019 (non-archival) & RepL4NLP 2020 [Paper]	
	Dohare, S., Gupta, V. , Karnick, H., <i>Unsupervised Semantic Abstractive Summarization</i> , ACL-SRW 2018 [Paper] [Poster]	
	Wadbude, R., Gupta, V. , Mekala, D., Karnick, H., <i>User Bias Removal in Review Score Prediction</i> , CODS-COMAD 2018 & DAB 2017 [Paper] [Poster]	
	Gupta, V.* , Mittal, S.*, Bhaumik, S., Roy, R. <i>Assisting Humans to Achieve Optimal Sleep by Changing Ambient Temperature</i> , BIBM 2016 , BHI 2016 & HI-DS 2016 [Paper] [Slides]	
	*represent equal contribution	
TECHNICAL MANUSCRIPTS	Gupta, V.* , Nokhiz, P.*, Dutta, C.*, Venkatasubramanian, S., <i>Equalizing Recourse Across Group</i> , ArXiv 2019 [PrePrint]	
	Mekala, D., Gupta, V. , Kar, P., Karnick, H., <i>Bayes-optimal Hierarchical Classification over Asymmetric Tree-Distance Loss</i> , Tech Report [PrePrint]	
	Mahajan, D., Gupta, V. , Keerthi, S., Sundararajan, S., <i>Efficient Estimation of Generalization Error and Bias-Variance Components of Ensembles</i> , Tech Report [PrePrint]	
MASTER THESIS	Product Categorization in E-Commerce using Distributional Semantics <i>Prof. Harish Karnick (IIT Kanpur) & Pradhuman Jhala (Flipkart.com)</i> [Thesis] [Slides]	
	<ul style="list-style-type: none"> Proposed a novel distributional semantics representation for text documents. Use the proposed representation with a novel two-level ensemble approach utilizing (with respect to the taxonomy tree) a path-wise, node-wise and depth-wise classifiers for product classification. 	
RESEARCH INTERNSHIPS	Bloomberg AI (IE and KG), London and India May 2022 - Sep 2022 Temporal Question Answering over Tables <i>Dr. Shuo Zhang, Dr. Ridho Reinanda</i>	
	Bloomberg AI (IE and KG), New York (Part-Time) Aug 2021 - Dec 2021 Information Extraction for Trustworthy Tabular Reasoning <i>Dr. Shuo Zhang, Dr. Alakananda Vempala, Dr. Yujie He, Dr. Temma Choi</i>	
	Bloomberg AI (IE and KG), New York May 2021 - Aug 2021 Information Extraction for Trustworthy Tabular Reasoning <i>Dr. Shuo Zhang, Dr. Alakananda Vempala, Dr. Yujie He, Dr. Temma Choi</i>	
	Facebook AI Research (Assistant), Seattle Sep 2020 - Dec 2020 Efficient and Effective Semantic Parsing <i>Dr. Denis Savenkov (Research Scientist)</i>	
	IBM Research, Thomas J. Watson Research Center May 2019 - Aug 2019 Contrastive Explanations for Natural Language Task <i>Dr. Kush R Varshney (Research Manager)</i>	
	Microsoft Research India, Bangalore May 2016 - Jul 2016 Estimation of Generalization Error for Ensembles <i>Dr. Sundararajan Sellamanickam (Principal Applied Scientist)</i>	
	Flipkart Internet Pvt. Ltd., Bangalore (Part Time) Aug 2015 - July 2016 Web Scale Product Classification <i>Pradhuman Jhala (Principal Architect)</i>	
	Flipkart Internet Pvt. Ltd., Bangalore May 2015 - July 2015 Web Scale Product Classification <i>Pradhuman Jhala (Principal Architect)</i>	
	Samsung R&D Institute, Bangalore May 2014 - July 2014 Mobile and Healthcare Solution Y2014 <i>Sandip Bhaumik (Group Manager) & Raj Roy (Manager)</i>	
	Synopsys Inc., Bangalore May 2013 - July 2013 DALI Verification System Coverage Visualization <i>Yogesh Pandey (Group Director)</i>	
Teaching Experience	Guest Instructor: for CS 2420 - Introduction to Algorithms and Data Structures, Spring 2022.	
	Taught two lecture on Graph Data Structures and Algorithms. Feedback available on request.	

Guest Instructor: for CS 2420 - Introduction to Algorithms and Data Structures, Fall 2021. Taught two lecture on Graph Data Structures and Algorithms. Design the weekly assignment and the practice lab session. Feedback available on request.

Workshop Organization: Co-Chaired to organized the NAACL-HLT Student Research Workshop 2021.

Seminar Coordinator: for CS 7941-Data Science Seminar (Spring 2021, Fall 2021-2020, Summer 2020) along with two University of Utah professors (Prof. Jeff Phillips and Prof. Aditya Bhaskara)

ACM iKDD Student Ninja: Helping ACM SIGKDD India Chapter (iKDD) to grow among the next-generation Data Science (DS) enthusiasts. Involved leadership roles in organizing iKDD activities as well as creating and running new periodic programs.

Teaching Mentor: for CS 6355 Spring 2021 Structured Prediction. Involves office hours for doubt clearing, assignment and examination evaluation.

Tutor: for CS 6150 Fall 2021 Advanced Algorithms. Involved guidance with the weekly assignments and the course material. Feedback available on request.

Teaching Mentor: for CS 6150 Fall 2019 Advanced Algorithms. Involves office hours for doubt clearing, assignment and examination evaluation.

Teaching Assistant: for MLT 2016 - Machine Learning Tool and Technique: Mentored a group of 30 M-Tech students part of a course on Machine Learning. Set up a labeling software for project work, resulting in a new dataset.

Teaching Assistant: for OS 2016 - Operating System: Mentored a group of 30 M-Tech students part of an introductory course on Operating System.

MISCELLANEOUS

[\[Professional Services\]](#), [\[Seminars Talks\]](#), [\[Research Grants\]](#), [\[Mentored Students\]](#)