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

Natural Language Inference, Representation Learning, Extreme Multi-Label Learning, Ethics in Artificial Intelligence

Current Position

School of Computing, University of Utah — Aug 2018 - Present
Graduate Research Assistant — Advisor: Prof. Vivek Srikumar

Education

School of Computing, University of Utah — Aug 2018 - Present
PhD, Computer Science, School of Computing
Advisor: Prof. Vivek Srikumar (CGPA: 3.9/4)

Indian Institute of Technology, Kanpur — July 2015 - May 2016
M. Tech, Computer Science and Engineering
Advisor: Prof. Harish Karnick (CGPA: 9.3/10)

Indian Institute of Technology, Kanpur — July 2011 - July 2015
B. Tech, Computer Science and Engineering (CGPA: 7.5/10)

Previous Employment

Microsoft Research Lab, India — Oct 2016 - Aug 2018
Research Fellow, Machine Learning and Natural Language Application — Dr. Nagarajan Natarajan,
Dr. Praneeth Netrapalli & Dr. Prateek Jain

Scholastic Achievements

One of the recipients of the Bloomberg Data Science Fellowship 2021-2022.

Selected to attend Virtual-HLF 2020, MLSS 2018 Madrid, MLSS 2020 London.

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

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**

Publications

Gupta, V., Zhang, S., Vempala, A., He, Y., Choji, T., Srikumar V., *Right for the Right Reason: Evidence Extraction for Trustworthy Tabular Reasoning*, ACL 2022 [Preprint]

Gupta, V., Bhat, R., Ghosal, A., Srivastava, M., Singh, M., Srikumar V., *Is My Model Using The Right Evidence? Systematic Probes for Examining Evidence-Based Tabular Reasoning*, TACL 2022 [Preprint]

Gupta, V., Shrivastava, A., Sagar, A., Aghajanyan, A., Savenkov. D., *RetroNLU: Retrieval Augmented Task Oriented Semantic Parsing*, under review [Preprint]

Jain, N., **Gupta, V.**, Rai, A., Kumar, G., *TabPert: An Effective Platform for Tabular Perturbation* EMNLP 2021, Demo track [Paper] [Project]

Gupta, V., Neeraja, J., 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 Page]

Gupta, A., **Gupta, V.**, *Unsupervised Contextualized Document Representation*, SustaiNLP 2021 at EMNLP 2021 [Paper]

Gupta, V., Saw A., Nokhiz, P., Gupta, H., Talukdar, P., *Improving Document Classification using Multi-Sense Embeddings*, NAACL-SRW 2019 (Oral) & ECAI 2020 (Oral) [Paper] [Blog]

Gupta, V., Bharti P., Nokhiz, P., Karnick, H., *SumPubMed: Summarization Dataset of PubMed Scientific Articles*, ACL-IJCNLP SRW 2021 [Preprint] [Dataset]

Gupta, V., Kumar, A., Nokhiz, P., Netrapalli, P., Rai, , P., Talulkdar, *P-SIF: Document Embeddings using Partition Averaging*, AAAI 2020 (Oral) [Paper] [Appendix] [PPT] [Poster] [Blog]

Gupta, V., Nokhiz, P., Dutta, C., Venkatasubramanian, S., *Equalizing Recourse Across Group*, ArXiv 2019 [PrePrint]

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*, AACL-IJCNLP 2020 [Paper]

Raunak, V., Dalmia, S., **Gupta, V.**, Metze, F.,*On Long-Tailed Phenomena in Neural Machine Translation*, EMNLP 2020 (Findings) & SPNLP 2020 [Paper]

Yadav, P., Yadav, P., Nokhiz, P., **Gupta, V.**, *Unbiasing Review Ratings with Tendency based Collaborative Filtering*, AACL-IJCNLP SRW 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 (Oral) [Paper] [Slides] [Poster]

Gupta, V., Saw A., Gupta, H., Nokhiz, P., Talukdar, P., *Word Polysemy Aware Document Vector Estimation*, NAACL-SRW 2019 (non-archival) [PrePrint]

Raunak, V., **Gupta, V.**, Metze, F., *Effective Dimensionality Reduction for Word Embeddings*, RepL4NLP 2019 (Oral) [Paper] [Poster]

Raunak, V., Kumar, V., **Gupta, V.**, Metze, F., *On Dimensional Linguistic Properties of the Word Embedding Space*, ACL-SRW 2019 (non-archival) & RepL4NLP 2020 (Oral) [Paper]

Dohare, S., **Gupta, V.**, Karnick, H., *Unsupervised Semantic Abstractive Summarization*, ACL-SRW 2018 [Paper] [Poster]

Mekala, D., **Gupta, V.**, Paranjape, B., Karnick, H. *Sparse Composite Document Vectors using soft clustering over distributional representations*, EMNLP 2017 (Oral) [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]

Wadbude, R., **Gupta, V.**, Mekala, D., Karnick, H., *User Bias Removal in Review Score Prediction*, CODS-COMAD 2018 & DAB 2017 (Oral) [Paper] [Poster]

Gupta, V., Mittal, S., Bhaumik, S., Roy, R. *Assisting Humans to Achieve Optimal Sleep by Changing Ambient Temperature*, BIBM 2016, BHI 2016 & HI-DS 2016 (Oral) [Paper] [Slides]

Mekala, Dheeraj., **Gupta, V.**, Kar, P., Karnick, H., *Bayes-optimal Hierarchical Classification over Asymmetric Tree-Distance Loss*, Tech Report [PrePrint]

Mahajan, D., **Gupta, V.**, Keerthi, S., Sundararjan, S., *Efficient Estimation of Generalization Error and Bias-Variance Components of Ensembles*, Tech Report [PrePrint]

Master Thesis

Product Categorization in E-Commerce using Distributional Semantics
Prof. Harish Karnick (IIT Kanpur) & Pradhuman Jhala (Flipkart.com) [Thesis] [Slides]

- 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), New York (Part-Time) Aug 2021 - Dec 2021

Information Extraction for Trustworthy Tabular Reasoning *Dr. Shuo Zhang, Dr. Alakananda Vempala, Dr. Yujie He, Dr. Temma Choi*

Bloomberg AI (IE and KG), New York May 2021 - Aug 2021
Information Extraction for Trustworthy Tabular Reasoning *Dr. Shuo Zhang, Dr. Alakananda Vempala, Dr. Yujie He, Dr. Temma Choi*

Facebook AI Research (Assistant), Seattle Sep 2020 - Dec 2020
Efficient and Effective Semantic Parsing *Dr. Denis Savenkov (Research Scientist)*

IBM Research, Thomas J. Watson Research Center May 2019 - Aug 2019
Contrastive Explanations for Natural Language Task *Dr. Kush R Varshney (Research Manager)*

Microsoft Research India, Bangalore May 2016 - Jul 2016
Estimation of Generalization Error for Ensembles *Dr. Sundararajan Sellamanickam (Principal Applied Scientist)*

Flipkart Internet Pvt. Ltd., Bangalore (Part Time) Aug 2015 - July 2016
Web Scale Product Classification *Pradhuman Jhala (Principal Architect)*

Flipkart Internet Pvt. Ltd., Bangalore May 2015 - July 2015
Web Scale Product Classification *Pradhuman Jhala (Principal Architect)*

Samsung R&D Institute, Bangalore May 2014 - July 2014
Mobile and Healthcare Solution Y2014 *Sandip Bhaumik (Group Manager) & Raj Roy (Manager)*

Synopsys Inc., Bangalore May 2013 - July 2013
DALI Verification System Coverage Visualization *Yogesh Pandey (Group Director)*

Teaching Experience

Guest Instructor: for CS 2420 - Introduction to Algorithms and Data Structures, Spring 2022. Taught two lecture on Graph Data Structures and Algorithms.

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.

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

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/Awards], [Mentored Students]