

# Vivek Gupta

Machine Learning & Natural Language Processing

CONTACT INFORMATION	The Preserve Apartment 543 South 900 East, Apt. C6, Salt Lake City, UT 84102 United States of America	<a href="https://github.com/vgupta123">vgupta123.github.io</a> <a href="mailto:keviv9@gmail.com">keviv9@gmail.com</a> +1-801-558-7272 <a href="#">Google Scholar</a>
RESEARCH INTERESTS	Natural Language Inference, Representation Learning, Extreme Multi-Label Learning, Ethics in Artificial Intelligence	
CURRENT POSITION	<b>School of Computing, University of Utah</b> <i>Graduate Research Assistant</i>	Aug 2018 - Present Advisor: Prof. Vivek Srikumar
EDUCATION	<b>School of Computing, University of Utah</b> PhD, Computer Science, School of Computing Advisor: Prof. Vivek Srikumar (CGPA: 3.9/4)	Aug 2018 - Present
	<b>Indian Institute of Technology, Kanpur</b> M. Tech, Computer Science and Engineering Advisor: Prof. Harish Karnick (CGPA: 9.3/10)	July 2015 - May 2016
	<b>Indian Institute of Technology, Kanpur</b> B. Tech, Computer Science and Engineering (CGPA: 7.5/10)	July 2011 - July 2015
EMPLOYMENT	<b>Microsoft Research Lab, India</b> <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	Selected among the top 7 finalist in <a href="#">Ericson's Innovation Awards, India</a> in 2016 Secured <b>All India Rank 183</b> in <b>IIT-JEE 2011</b> amongst more than 500,000 candidates Qualified for <b>Kishore Vaigyanik Protsahan Yojana (KVPY) scholarship, 2011</b> Selected among top 1% of students, <b>Indian National Physics Olympiads, 2011</b> Selected among top 1% of students, <b>Indian National Chemistry Olympiads, 2011</b>	
PUBLICATIONS	<b>Gupta, V.</b> , Neeraja, J., Srikumar, V. <i>Incorporating External Knowledge to Enhance Tabular Reasoning</i> , <a href="#">NAACL 2021</a> <a href="#">[Paper]</a> <a href="#">[Project]</a> <b>Gupta, V.</b> , Mehta, M., Nokhiz, P., Srikumar, V. <i>InfoTabS: Inference on Tables as Semi-structured Data</i> , <a href="#">ACL 2020</a> <a href="#">[Paper]</a> <a href="#">[Project Page]</a> <b>Gupta, V.</b> , Saw A., Nokhiz, P., Gupta, H., Talukdar, P., <i>Improving Document Classification using Multi-Sense Embeddings</i> , <a href="#">ECAI 2020</a> (Oral) <a href="#">[Paper]</a> <a href="#">[Blog]</a> <b>Gupta, V.</b> , Bharti P., Nokhiz, P., Karnick, H., <i>SumPubMed: Summarization Dataset of PubMed Scientific Articles</i> , <a href="#">ACL-IJCNLP SRW 2021</a> <a href="#">[Preprint]</a> <a href="#">[Dataset]</a> <b>Gupta, V.</b> , Kumar, A., Nokhiz, P., Netrapalli, P., Rai, P., Talulkdar, <i>P-SIF: Document Embeddings using Partition Averaging</i> , <a href="#">AAAI 2020</a> (Oral) <a href="#">[Paper]</a> <a href="#">[Appendix]</a> <a href="#">[PPT]</a> <a href="#">[Poster]</a> <a href="#">[Blog]</a> <b>Gupta, V.</b> , Nokhiz, P., Dutta, C., Venkatasubramanian, S., <i>Equalizing Recourse Across Group</i> , ArXiv 2019 <a href="#">[PrePrint]</a> Uppal S., <b>Gupta, V.</b> , Swaminathan A., Zhang H., Mahata D., Gosangi R., Shah. R., Stent A., <i>Two-Stage Classification using Recasted Data for Low Resource Settings</i> , <a href="#">AAACL-IJCNLP 2020</a> <a href="#">[Paper]</a> Raunak, V., Dalmia, S., <b>Gupta, V.</b> , Metze, F., <i>On Long-Tailed Phenomena in Neural Machine Translation</i> , <a href="#">EMNLP 2020</a> (Findings) & <a href="#">SPNLP 2020</a> <a href="#">[Paper]</a>	

- Yadav, P., Yadav, P., Nokhiz, P., **Gupta, V.**, *Unbiasing Review Ratings with Tendency based Collaborative Filtering*, [AAACL-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., Sundararajan, 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.
  - Proposed a 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

- 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** 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)*

## MISCELLANEOUS

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