CONTACT The Preserve Apartment vgupta123.github.io
INFORMATION 543 South 900 East, Apt. C6,
Salt Lake City, UT 84102 +1-801-558-7272
United States of America Google Scholar

RESEARCH Natural Language Inference, Representation Learning, Extreme Multi-Label Learning, Ethics in Arti-INTERESTS ficial Intelligence

CURRENT School of Computing, University of Utah Aug 2018 - Present Position 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)

EMPLOYMENT Microsoft Research Lab, India Oct 2016 - Aug 2018

Research Fellow, Machine Learning and Natural Language Application Dr. Nagarajan Natararjan,

Dr. Praneeth Netrapalli & Dr. Prateek Jain

Scholastic Selected among the top 7 finalist in Ericson's Innovation Awards, India in 2016

ACHIEVEMENTS 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., Mehta, M., Nokhiz, P., Srikumar, V. InfoTabS: Inference on Tables as Semi-structured

Data, ACL 2020 [PrePrint] [Project Page]

Gupta, V., Saw A., Nokhiz, P., Gupta, H., Talukdar, P., Improving Document Classification using Multi-Sense Embeddings, ECAI 2020 (Oral) [PrePrint] [Blog]

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, submitted to FAccT 2021 [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 [PrePrint]

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

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) [PrePrint]

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.
- 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 Applied Scientist)

Dr. Sundararajan Sellamanickam (Principal

Flipkart Internet Pvt. Ltd., Bangalore

May 2015 - July 2015 Pradhuman Jhala (Principal Architect)

Web Scale Product Classification

Mars 2014 Inles 2014

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]