Tushaar Gangavarapu

Department of Computer Science
Cornell University, Ithaca, NY

Web: https://tushaargvs.github.io/
Email: tg352@cornell.edu

A. Education

2022−2024	Master of Science in Computer Science.
	Project (2023-24): *Structured state space models for language modeling.*
	Committee: Alexandar Sasha Rush (chair) and David Bindel.
	Project (2022-23): *Predictive modeling to identify conversational trajectories.*
	Advisors: Cristian Danescu-Niculescu-Mizil and Lillian Lee.
2015−2019	Undergraduate study in Information Technology with specialization in Artificial Intelligence (rank: 3/104), National Institute of Technology Karnataka.
	Thesis: *Psychological aspects and behavioral traits in social media language.*
	Committee: Ram Mohana Reddy Guddeti (advisor/chair), Sowmya Kamath S., Nagamma Patil, Biju R. Mohan.

Scholarships and honors

Huawei national scholarship for academic excellence, 2018−2019.

National higher secondary education scholarship for excellence (national rank: 10), 2013−2015.

South Indian Mathematics Olympiad medal of excellence (rank: 32), 2012.

B. Employment

01/2023−05/2023	Graduate Teaching Assistant, Language and Information (CS 4300; INFO 4300), Cornell University.
08/2022−01/2023	Graduate Teaching Assistant, Natural Language Processing (LING4744; COGSTAT4740; CS4/5740), Cornell University.
09/2021−08/2022	Applied Scientist, Kindle Content Quality Algorithms, Amazon.com, Inc.
01/2020−08/2021	Research Engineer (natural language processing), Kindle, Amazon.com, Inc.
06/2019−12/2019	Software Development Engineer, Consumer Engagement, Amazon.com, Inc.
2018−2019	Scientific researcher, Healthcare Analytics and Language Engineering lab, National Institute of Technology Karnataka.
05/2018−07/2018	Software Development Intern, Kindle Create, Amazon.com, Inc.
2017−2019	Research candidate, Human Centered Computing Group, National Institute of Technology Karnataka.
2017	Research Intern, Center for Pattern Recognition and Machine Intelligence, People's Education Society (PES) University.

C. Publications

Scholar: https://scholar.google.com/citations?user=C7v_cA8AAAAJ.

Research interests

Natural language processing and understanding.

Healthcare analytics and clinical decision-support systems.

Computational social science.

Journal articles

Tushaar Gangavarapu, Gokul S. Krishnan, Sowmya Kamath S., and Jayakumar Jeganathan. FarSight: Long-Term Disease Prediction Using Unstructured Clinical Nursing Notes. *IEEE Transactions on Emerging Topics in Computing*, 9(3):1151–1169, 2021. doi: 10.1109/TETC.2020.2975251.

Mayya V., Sowmya Kamath S., Gokul S. Krishnan, and Tushaar Gangavarapu. Multi-channel, convolutional attention based neural model for automated diagnostic coding of unstructured patient discharge summaries. *Future Generation Computer Systems*, 118:374–391, 2021. doi: 10.1016/j.future.2021.01.013.

Tushaar Gangavarapu, C. D. Jaidhar, and Bhabesh Chanduka. Applicability of machine learning in spam and phishing email filtering: review and approaches. *Artificial Intelligence Review*, 53(7):5019–5081, 2020a. doi: 10.1007/s10462-020-09814-9. [Review article].

Tushaar Gangavarapu, Aditya Jayasimha, Gokul S. Krishnan, and Sowmya Kamath S. Predicting ICD-9 code groups with fuzzy similarity based supervised multi-label classification of unstructured clinical nursing notes. *Knowledge-Based Systems*, 190:105321, 2020b. doi: 10.1016/j.knosys.2019.105321.

Tushaar Gangavarapu and Nagamma Patil. A novel filter–wrapper hybrid greedy ensemble approach optimized using the genetic algorithm to reduce the dimensionality of high-dimensional biomedical datasets. *Applied Soft Computing*, 81: 105538, 2019. doi: 10.1016/j.asoc.2019.105538.

Refereed full papers in conference proceedings

Yann Hicke, Abhishek Masand, Wentao Guo, and Tushaar Gangavarapu. Assessing the efficacy of large language models in generating accurate teacher responses. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*. Association for Computational Linguistics, 2023. to appear.

Tushaar Gangavarapu and Sriraghavendra Ramaswamy. Alexa, stop reading the references: Enhancing the reading experience in Kindle eBooks. In *Amazon Machine Learning Conference – AMLC 2022*. Amazon, 2022a.

Tushaar Gangavarapu and Sriraghavendra Ramaswamy. A figure is worth a thousand words, but where are the words?: Enhancing image experience in Kindle eBooks. In *Amazon Machine Learning Conference – AMLC 2022*. Amazon, 2022b.

Tushaar Gangavarapu and C. D. Jaidhar. A Novel Bio-inspired Hybrid Metaheuristic for Unsolicited Bulk Email Detection. In *Computational Science – ICCS 2020*, pages 240–254. Springer, 2020. doi: 10.1007/978-3-030-50420-5_18.

Aditya Jayasimha, Tushaar Gangavarapu, Sowmya Kamath S., and Gokul S. Krishnan. Deep Neural Learning for Automated Diagnostic Code Group Prediction Using Unstructured Nursing Notes. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, CoDS COMAD 2020, pages 152–160. Association for Computing Machinery, 2020. doi: 10.1145/3371158.3371176.

Tushaar Gangavarapu, Gokul S. Krishnan, and Sowmya Kamath S. Coherence-based Modeling of Clinical Concepts Inferred from Heterogeneous Clinical Notes for ICU Patient Risk Stratification. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 1012–1022, 2019b. doi: 10.18653/v1/K19-1095.

Tushaar Gangavarapu, Aditya Jayasimha, Gokul S. Krishnan, and Sowmya Kamath S. TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes. In *Natural Language Processing and Information Systems*, pages 195–207. Springer, 2019a. doi: 10.1007/978-3-030-23281-8_16.

Tushaar Gangavarapu, Himadri Pal, Pratyush Prakash, Suraj Hegde, and V. Geetha. Parallel OpenMP and CUDA Implementations of the N-Body Problem. In *Computational Science and Its Applications – ICCSA 2019*, pages 193–208. Springer, 2019c. doi: 10.1007/978-3-030-24289-3_16.

Bhabesh Chanduka, Tushaar Gangavarapu, and C. D. Jaidhar. A Single Program Multiple Data Algorithm for Feature Selection. In *Intelligent Systems Design and Applications – ISDA 2018*, pages 662–672. Springer, 2018. doi: 10.1007/978-3-030-16657-1_62.

D. Professional service

Recent invited talks

Learning to Predict: Tree-based Classification. Guest lecture at the Machine Learning University, Amazon.com, Inc., 2020.

Cognitive and Affective Assessments in Game-based Simulated Environments. Invited talk at the Department of Information Technology, National Institute of Technology Karnataka, 2020.

Greedy Evolutionary Feature Selection for Biomedical Data. Invited talk at the Department of Information Technology, National Institute of Technology Karnataka, 2020.

On the Convergence of High Performance Computing and Machine Intelligence. Invited talk at the National Workshop on High Performance Computing and Applications (HPCA 2019), National Institute of Technology Karnataka.

Exploring Latent Human Traits Through Social Media Modeling. Guest lecture at the Department of Information Technology, National Institute of Technology Karnataka, 2019.

Game-based Learning and Assessment: A Case Study of a Mobile-VR Game. Guest lecture at the Department of Information Technology, National Institute of Technology Karnataka, 2019.

Building Predictive Applications Using Social Media Digital Footprints. Invited talk at the National Workshop on Predictive Analytics and Applications (PAA 2019), National Institute of Technology Karnataka.

Refereeing

2022	Amazon Machine Learning Conference (AMLC) 2022: Healthcare Informatics and Natural Language Processing track.
2020	Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2020, Healthcare Analytics and Language Processing (HeAL) 2020 program co-chair.

Scientific activities

2020	Co-organizer of AQuA-tic ConScience: Books machine learning ideation and brainstorming group, Amazon.com, Inc.
2018	Research Assistant for Applied Soft Computing and Genetic Algorithms (IT355), National Institute of Technology Karnataka.
2018	Teaching Assistant for Computer Communications and Networking (IT251), National Institute of Technology Karnataka.
2016	Teaching Assistant for Machine Intelligence and Neural Learning Algorithms, Winter Mentorship Programme, National Institute of Technology Karnataka.

University committees

2022−present	Student member, Machine Learning in Medicine (MLIM), Cornell University (Ithaca) and Weill Cornell Medicine.
2016−2019	Executive member, Institution of Engineers (IE Code SIG), National Institute of Technology Karnataka.
2016−2019	Executive member, Web Enthusiasts' Club (Intelligence SIG), National Institute of Technology Karnataka.
2016−2019	National Institute of Technology Karnataka Student Ambassador, Intel Artificial Intelligence.
2016−2018	National Institute of Technology Karnataka Student Ambassador, CodeNation.

[Last updated: 08/15/2023]