Aman Khullar

https://amankhullar.github.io

akhullar8@gatech.edu

EDUCATION

Georgia Institute of Technology — Atlanta, GA
Ph.D. in Computer Science, HCI Specialization — *Aug. 2023 – Present*
Advisor: Dr. Neha Kumar

Georgia Institute of Technology — Atlanta, GA
Master's in Computer Science, ML Specialization (GPA: 4.00/4.00) — *Aug. 2021 – May 2023*
Coursework: Graduate Algorithms, Machine Learning, NLP, AI Storytelling

Netaji Subhas Institute of Technology, University of Delhi — New Delhi, India
Bachelor of Engineering in Information Technology — *Aug. 2015 – May 2019*
First Class with Distinction (CGPA: 8.65/10.00)

EXPERIENCE

Amazon — Virtual, GA
Applied Scientist Intern, Alexa Hybrid Science — *Aug. 2022 – Dec. 2022*

- Improved the training recipe to incorporate synthetic data to fine-tune speech recognition models
- Quantitatively evaluated role of synthetic speech and loss modification on speech recognition accuracy
- Compiled an initial paper draft of our work reporting speech recognition accuracy improvement of 2%
- Presented the paper, "Robust Speech Recognition via Large-Scale Weak Supervision" in paper discussion

Apple — Seattle, WA
AI/ML Intern, Siri Text-To-Speech — *May 2022 – Aug. 2022*

- Improved end-to-end speech synthesis with diffusion models using advanced encoder architectures

Georgia Institute of Technology — Atlanta, GA
Graduate Research Assistant, T+ID lab (advisor: Prof. Michael Best) — *Jan. 2022 – May 2022*

- Developed a contextual entity replacement pipeline to generate synthetic training data in low resource languages — taking cues from hate speech data in English — to facilitate training of hate speech detection classifiers
- Our developed methodology outperformed the classifiers trained on machine translated hate speech data from English to Hindi — the only other alternative — by a factor of 1.4% in accuracy and 2.2% in F1 score

Gram Vaani — Gurugram, India
Research Associate (advisor: Prof. Aaditeshwar Seth) — *Mar. 2020 – Jul. 2021*

- Developed automatic content moderation tools for in-house moderators, voice based surveys for user data collection, automatic question-answering system for IVR platforms — deployed in a product used by 800,000 rural households
- Conducted quantitative and qualitative evaluation of the effect of introducing AI for automating Interactive Voice Response (IVR) based voice applications — published and presented the results in 3 peer-reviewed publications
- Developed MLOps pipeline for continuous model training — triggered on accuracy of ML model drifting below 95%
- Awarded the Gram Vaani Star Performer award for two consecutive quarters

Product Labs, IIIT Hyderabad — Hyderabad, India
Research Engineer (advisors: Prof. Manish Shrivastava and Mr. Prakash Yalla) — *Jul. 2019 – Feb. 2020*

- Developed a deep learning based document question-answering model to extract answers from a group of documents; achieved 3 seconds response time and 90% accuracy over 137 page document — evaluated by Indian Army
- Productized the research — model formed core product for Subtl.ai's first instance shipped to 2,500 users

TEACHING EXPERIENCE

- **Teaching Assistant**: CS7650 Natural Language Processing (Spring 2023 under Prof. Alan Ritter)

Selected Publications

- **A. Khullar***, D. Nkemelu*, et al. 2023. Hate Speech Detection in Limited Data Contexts using Synthetic Data Generation. *ACM Journal on Computing and Sustainable Societies* [pdf, talk].

- **A. Khullar**, et al. 2021. Costs and Benefits of Conducting Voice-based Surveys Versus Keypress-based Surveys on Interactive Voice Response Systems. *ACM COMPASS* [pdf, talk].

- **A. Khullar**, et al. 2021. Experiences with the Introduction of AI-based Tools for Moderation Automation of Voice-based Participatory Media Forums. *India HCI* and *KDD Workshop on Data Science for Social Good* [pdf].

- **A. Khullar**, et al. 2021. Early Results from Automating Voice-based Question-Answering Services Among Low-income Populations in India. *ACM COMPASS* [pdf, talk].

- M. Johri, S. Agarwal, **A. Khullar**, D. Chandra, V. S. Pratap and A. Seth. 2021. The first 100 days: How has COVID-19 affected poor and vulnerable groups in India? *Health Promotion International* [pdf, talk, policy brief].

- **A. Khullar*** and U. Arora*. 2020. MAST: Multimodal abstractive summarization with trimodal hierarchical attention. *EMNLP Workshop on NLP Beyond Text* [pdf].

Projects

- **Knowledge Graphs + Language Models**: Combined a knowledge graph with the GPT-J model to reduce inconsistencies in generated text by rejecting contradicting sentences using NLI and enhanced prompting [video]

- **Digital Scribes**: Worked with Stanford Healthcare AI Applied Research team to understand the gaps in current digital scribes technology and how we can improve upon them – consulted the Doctors in resident from an NLP standpoint

- **HUMAN: Hate-speech Understanding by Multimodal Attention Network**: Explained the predictions MultiModal BiTransformer Model (Kiela et al.), improved model accuracy by 1.2% using image autoencoder [link]

- **Autonomous Mars Rover**: Co-founder and team lead of ARES, NSIT's first Mars rover team — ranked among top 71 and 30 teams in University and European Rover Challenges respectively

- **MultiModal BiDirectional Attention Flow model**: Augmented a machine comprehension model (BiDAF) to be used for multimodal summarization as part of my Bachelor's thesis — awarded 3rd highest grade in the batch

- **Interaction of Game Theory and Machine Learning in a connected World**: Presented a survey report highlighting intersection of these fields in electronic market design and selfish routing — awarded highest grade in batch

Awards and Honors

- **Computing For Good Fellowship**: Among two students selected for C4G fellowship at Georgia Tech for Spring 2023

- **AllenNLP Hacks 2021**: Impact Award runner-up to help identify communities subjected to hate speech

- **Google Research India AI Summer School**: Among 150 candidates selected in India for AI summer school 2020

- **Harvard College Project for Asian and International Relations**: Among 300 delegates selected for HPAIR 2019

- **India's Brightest 33 Engineers**: Among the cohort of 33 students selected by the Economic Times in 2018 [talk]

- **Global Food Ambassador**: Among 278 ambassadors worldwide, selected by Thought For Food organization, in 2017, to work with food entrepreneurs and find innovative solutions to the problem of feeding 9 billion people by 2050

Academic Service

- **Reviewer**: AfriCHI, 2023
- **Reviewer**: CHI, 2023
- **Reviewer**: IET Computer Vision Journal, 2022
- **Program Committee**: NLP Beyond Text Workshop, The Web Conference, 2021
- **Shadow Program Committee**: ACM COMPASS, 2021
- **Program Committee**: NLP Beyond Text Workshop, EMNLP, 2020

Skills

- **Languages and Frameworks**: Python, C/C++, PyTorch, TensorFlow, Git, MySQL, Django

---

*Equal contribution