Sonal Kumar

+1 240 501 3161 sonalkum@umd.edu realsonalkumar sonalkum sonalkum.github.io

EDUCATION

University of Maryland, College Park **2023 – Present**
PhD in Computer Science, Advised by ***Prof. Dinesh Manocha & Prof. Ramani Duraiswami*** *- 3.8/4* *Maryland, USA*

CHRIST (Deemed to be University), Bangalore **2017 – 2021**
Bachelor of Technology in Computer Science and Engineering - ***9.7****/10* *Bangalore, India*

Publications

Synthio: Augmenting Small-Scale Audio Classification Datasets with Synthetic Data
Sreyan Ghosh*, ***Sonal Kumar****, Zhifeng Kong, Rafael Valle, Bryan Catanzaro, Dinesh Manocha
ICLR 2025

MMAU: A Massive Multi-Task Audio Understanding and Reasoning Benchmark
S Sakshi*, Utkarsh Tyagi*, ***Sonal Kumar****, Ashish Seth*, Ramaneswaran Selvakumar, Oriol Nieto, Ramani Duraiswami, Sreyan Ghosh*, Dinesh Manocha
ICLR 2025

Visual Description Grounding Reduces Hallucinations and Boosts Reasoning in LVLMs
Sreyan Ghosh*, Chandra Kiran Reddy Evuru*, ***Sonal Kumar****, Utkarsh Tyagi, Oriol Nieto, Zeyu Jin, Dinesh Manocha
ICLR 2025

GAMA: A Large Audio-Language Model with Advanced Audio Understanding and Complex Reasoning Abilities
Sreyan Ghosh*, ***Sonal Kumar****, Ashish Seth, Chandra Kiran Reddy Evuru, Utkarsh Tyagi, S Sakshi, Oriol Nieto, Ramani Duraiswami, Dinesh Manocha
EMNLP 2024

EH-MAM: Easy-to-Hard Masked Acoustic Modeling for Self-Supervised Speech Representation Learning
Ashish Seth, Ramaneswaran S, S Sakshi, ***Sonal Kumar***, Sreyan Ghosh, Dinesh Manocha
EMNLP 2024

A Closer Look at the Limitations of Instruction Tuning
Sreyan Ghosh*, C. K. Evuru*, ***Sonal Kumar****, Ramaneswaran S, D. Aneja, Z. Jin, R. Duraiswami, Dinesh Manocha
ICML 2024

LipGER: Visually-Conditioned Generative Error Correction for Robust Automatic Speech Recognition
Sreyan Ghosh, ***Sonal Kumar***, Ashish Seth, Purva Chiniya, Utkarsh Tyagi, Ramani Duraiswami, Dinesh Manocha
Interspeech 2024

AV-RIR: Audio-Visual Room Impulse Response Estimation
Anton Ratnarajah, Sreyan Ghosh, ***Sonal Kumar***, Purva Chiniya, Dinesh Manocha
CVPR 2024

CompA: Addressing the Gap in Compositional Reasoning in Audio-Language Models
Sreyan Ghosh*, Ashish Seth*, ***Sonal Kumar****, Utkarsh Tyagi, C. K. Evuru, Oriol Nieto, Dinesh Manocha
ICLR 2024

RECAP: Retrieval-Augmented Audio Captioning
Sreyan Ghosh*, ***Sonal Kumar****, Chandra Kiran Reddy Evuru, Ramani Duraiswami, Dinesh Manocha
ICASSP 2024

Span Classification with Structured Information for Disfluency Detection in Spoken Utterances
Sreyan Ghosh, ***Sonal Kumar***, Yaman Kumar Singla, Rajiv Ratn Shah, S. Umesh
Interspeech 2022

Do Vision-Language Models Understand Compound Nouns?
Sonal Kumar*, Sreyan Ghosh*, S Sakshi, Utkarsh Tyagi, Dinesh Manocha
NAACL 2024

ASPIRE: Language-Guided Augmentation for Robust Image Classification
Sreyan Ghosh*, C. K. Evuru*, ***Sonal Kumar***, S. Sakshi, Utkarsh Tyagi, Dinesh Manocha
ACL 2024

ABEX: Data Augmentation for Low-Resource NLU via Expanding Abstract Descriptions

Sreyan Ghosh, Utkarsh Tyagi, ***Sonal Kumar***, Chandra Kiran Reddy Evuru, Ramaneswaran S, S Sakshi, Dinesh Manocha
ACL 2024

CoDa: Constrained Generation based Data Augmentation for Low-Resource NLP
Chandra Kiran Reddy Evuru*, Sreyan Ghosh*, ***Sonal Kumar***, Ramaneswaran S, Utkarsh Tyagi, Dinesh Manocha
NAACL 2024

DALE: Generative Data Augmentation for Legal NLP
Sreyan Ghosh*[1], C. K. Evuru*, ***Sonal Kumar***, S. Sakshi, Utkarsh Tyagi, Dinesh Manocha
EMNLP 2023

CoSyn: Detecting Implicit Hate Speech in Online Conversations Using a Context Synergized Hyperbolic Network
Sreyan Ghosh, Manan Suri, Purva Chiniya, Utkarsh Tyagi, ***Sonal Kumar***, Dinesh Manocha
EMNLP 2023

ACLM: A Selective-Denoising based Generative Data Augmentation Approach for Low-Resource Complex NER
Sreyan Ghosh*, Utkarsh Tyagi*, Manan Suri, *Sonal Kumar*, S Ramaneswaran, Dinesh Manocha
ACL 2023

BioAug: Conditional Generation based Data Augmentation for Low-Resource Biomedical NER
Sreyan Ghosh*, Utkarsh Tyagi*, ***Sonal Kumar****, Dinesh Manocha
SIGIR 2023

RESEARCH EXPERIENCE

GAMMA Lab, University of Maryland — **College Park, Maryland**
Research Assistant — *October 2022 – Present*

- As a part of GAMMA Lab at UMD, worked on Data Augmentations for low resource NLP problems such as Complex NER, Bio-Medical NER and Legal NER. Also working on Audio and Speech Processing along with Audio-Language research. Published at ICML, ICLR, EMNLP, ICASSP, ACL, NAACL, SIGIR, Interspeech.
- Advised by Prof. Dinesh Manocha

PIRL Lab, University of Maryland — **College Park, Maryland**
Research Assistant — *August 2023 – Present*

- As a part of PIRL Lab at UMD, I am working on decoding imagined speech information from human brain's MEG signals and Audio and Speech Processing.
- Advised by Prof. Ramani Duraiswami

INDUSTRY EXPERIENCE

Adobe Inc. — **San Francisco, USA**
Research Scientist Intern — *May 2024 – November 2024*

- Working as a research scientist intern at Adobe in the audio research team on Audio Generative Models.

Cisco Systems — **Bangalore, India**
Software Engineer 2 — *August 2021 – August 2023*

- Responsible for automating workflow to perform bulk OS upgraded for in-production network devices with minimal impact. Led to saving of 3-5 man hours per upgrade.
- Monitored and led successful upgrade of 1000+ in production network devices.
- Led a team of 5 to build chat bot which leveraged internal data sources to solve queries related to Customer Delivery using Machine Learning techniques.
- Built Network Ready For Use (NRFU) tool to automate testing of newly deployed network saving 5-6 man hours per testing.

Software Engineer Intern — *Jan 2021 – July 2021*

- Worked with the Data Center team to develop automated workflows to test in production network devices.
- Received appreciation for the excellent performance in Intern Spotlight 2021

IHS Markit — **Bangalore, India**
Software Engineer Intern — *June 2020 – July 2020*

- Worked on NLP and developed a PoC for NER system to extract information from PDF documents
- Kapow/Kofax RPA for automating workflow, and dashboard creation.

[1] *Denotes co-first authors

TECHNICAL SKILLS

Skills: Deep Learning, Speech Processing, NLP, Computer Vision, Multimodal Learning, Generative AI, Research and Development (R&D)
Languages: (Highly Proficient) Python, JavaScript, C++, Java, SQL
Frameworks: Pytorch, FastAI, Tensorflow, Scikit-Learn, NLTK, Flask

PROJECTS

MoM.ai | *Deep Learning, Natural Language Processing*

- Built an AI enabled web app which takes in a pre-recorded audio, performs real time speaker diarization, generates a transcript with timestamps and placeholder for manual speaker tagging, and generates Minutes of the meeting (MoM).

SpeeQL | *Deep Learning, Natural Language Processing*

- Built a web-app to enhance customer experience in store. Used Speech to Text and Named Entity Recognition to recognise products and brands, and used OCR to extract text from the image of list of products.

ACHIEVEMENTS & EXTRA CURRICULAR

- **Cisco Systems, March 2023** - *Recognised for outstanding performance and contribution to AI research*
- **Mentor2Go, Jan 2022 - Aug 2022** - *Mentor*
- **P&G Global Innovation Challenge, May 2020** - *Best Idea Prize*
- **CHRIST (Deemed to be University), Feb 2020** - *CHRISTITE Spirit Award* for bringing recognition to the university.
- **Neuron - AI Club, CHRIST (Deemed to be University), June 2019 - Aug 2019** - *President*
- **Hindustan Unilever BFS Technology Hackathon, Sep 2019** - *Winner*
- **Kaggle Quora Insincere Questions Classification, Feb 2019** - *Bronze Medal*
- **MLDS-2019 Republic Day Hackathon,TEG Analytics, Jan 2019** - *Winner*