Sreyan Ghosh

+1 (240) 899-7685   sreyang@umd.edu

- LinkedIn   • GitHub   • Google Scholar   • Portfolio

EDUCATION

Bachelor of Technology in Computer Science, Christ University, GPA: 8.69/10.00   **April 2016 — April 2020**

Master's in Computer Science, University of Maryland, College Park, GPA: -/-   **August 2022 — April 2024 (Expected)**
Advisor: Dr. Dinesh Manocha

PROFESSIONAL EXPERIENCE

Deep Learning Solutions Architect (Natural Language and Speech Processing)   **April 2022 — August 2022**
NVIDIA   Bangalore, India

- Worked as a senior solutions architect in the professional services team at NVIDIA. Responsible for delivering deep-learning based NLP solutions to NVIDIA's premier customers around the globe.
- Contributed to AI R&D at NVIDIA. Published 2 papers at **IEEE SLT 2022**.

Software Engineer II   **August 2020 — March 2022**
Cisco Systems   Bangalore, India

- Worked as a senior software engineer in the automation and orchestration team under the Customer Experience BU. Built network assurance solutions for Cisco's telecom customers, leveraging state-of-the-art algorithms for anomaly detection at scale. Built a critical component in Cisco's first telemetry-based network assurance solution.
- Lead the development of an AI-based network security system for one of Cisco's telecom customers.
- Was part of the AI team that developed Cisco's first contact center solution, leveraging state-of-the-art NLP algorithms.
- Contributed to AI R&D at Cisco by representing Cisco at various conferences.

RESEARCH AND TEACHING EXPERIENCE

Teaching Assistant   **Fall 2022**
University of Maryland, College Park   College Park, Maryland, U.S.A.

- TA for Introduction to NLP (CMSC 470) for Fall 2022.

Research Assistant   **Fall 2022**
Gamma Lab @ UMCP   College Park, Maryland, U.S.A.

- Conducting research on audio-visual speech enhancement and SLP advised by Dr. Dinesh Manocha.

Project Associate (Research) [Link]   **June 2021 — August 2022**
Speech Lab, Department of Electrical Engineering, Indian Institute of Technology Madras   Chennai, India

- Worked under the supervision of **Dr. S. Umesh** in the area of self-supervised learning for Speech and Audio processing. Exploring techniques to devise lighter-weight models and efficient algorithms to make supervised and self-supervised learning in speech and audio amenable to resource-constrained scenarios (both data and compute). Paper accepted to **SAS Workshop at AAAI 2022**, **Interspeech 2022**, **IEEE JSTSP Special Issue** and **IEEE SLT 2022**.

Research Assistant (part-time) [Link]   **December 2019 — August 2022**
MIDAS Labs, IIIT-Delhi   Delhi, India

- Worked under the supervision of **Dr. Rajiv Ratn Shah** in the areas of Speech and Natural Language Processing. Worked on building ASR systems for low-resource Indian Languages (mono and multi-lingual) and Indian Accented English which served as a critical component for other systems built by the lab. Worked in the field of content moderation in modalities of both speech and text. Currently exploring multi-modal techniques for identifying disfluencies in spoken utterances. Published papers at **AAAI 2021, ACL 2021** and **Interspeech 2022**.

Research Assistant & Teaching Assistant   **December 2018 — December 2019**
Christ University   Bangalore, India

- Worked under the supervision of **Dr. Samiksha Shukla** and **Dr. Aynur Unal** on detecting Bipolar Disorder in individuals from clinical records. Used multi-variate time-series modelling to capture the sequential aspect of medical records recorded on weekly intervals. Worked was done in collaboration with Stanford University. Paper accepted to Springer Singapore. Also presented our work at Yale Global Health Innovation Conference.
- Worked under **Prof. Julian Benedict** on a project based on AI based pre-fetching for web caching.
- Taught students basic to advanced concepts in AI, including conducting practical workshops and basics of DSA.

Internships

Google Summer of Code   **April 2022 — August 2022**
Tensorflow, Google   Remote
- Working on building deep learning based NLP (speech and text) notebooks using Tensorflow and Keras.
- Link to PRs and code contributed on personal website.

Research Intern   **January 2020 — June 2020**
MIDAS Labs, IIIT-Delhi   Delhi, India
- Worked on a project, End-to-End Named Entity Recognition from English Speech under the guidance of **Dr. Rajiv Ratn Shah** as part of my bachelors thesis. Paper accepted at **Interspeech 2020**.

Technology Intern   **January 2020 — June 2020**
Cisco Systems   Bangalore, India
- Worked on building a VOIP (Voice Over IP) Traffic Analyzer to detect anomalous SIP messages using machine learning.

Data Science Intern   **December 2019 — January 2020**
Noodle.ai   Bangalore, India
- Worked on multivariate time-series anomaly detection in high-frequency IoT sensor data obtained from steel manufacturing machines.

Data Science Intern   **April 2019 — May 2019**
TEG Analytics   Bangalore, India
- Worked under the healthcare intelligence division to provide insights from insurance plan enrollment data, for private insurance companies in the US.
- Used Machine Learning and Deep Learning techniques to predict plan enrollment for insurance companies.

Skills

Programming	Python, Java, C, MySQL, HTML, CSS
Frameworks	PyTorch, Keras, Tensorflow, Django, Flask, Spark
Tools	GIT, Android, Tableau, Power BI, AWS, GCP, Rest API, Docker, K8s
Concepts	Speech and Natural Language Processing, Software Development, Functional programming, Object- oriented programming, Machine Learning, Deep Learning, Image Processing, Cloud Computing

Publications (Accepted & Pending) and Patents

1. **Ghosh, S.**, Seth, A. & Umesh, S. Decorrelating Feature Spaces for Learning General Purpose Audio Representations. **Link**. *IEEE JSTSP Special Issue on Self-Supervised Learning for Speech and Audio Processing.*

2. Prasad, L. V. S. V. D., **Ghosh**, **S.** & Umesh, S. CCC-Wav2Vec 2.0: Clustering Aided Cross Contrastive Self-Supervised Learning of Speech Representations. *IEEE SLT 2022.*

3. Prasad, L. V. S. V. D., **Ghosh, S.** & Umesh, S. PADA: Pruning Assisted Domain Adaptation for Self-Supervised Speech Representations. **Link**. *IEEE SLT 2022.*

4. **Ghosh, S.**, Lepcha, S., Sakshi, S. & Shah, R. R. DeToxy: A Large-Scale Multimodal Dataset for Toxicity Classification in Spoken Utterances. **Link**. *Interspeech 2022.*

5. **Ghosh, S.**, Seth, A. & Umesh, S. Span Classification with Structured Information for Disfluency Detection in Spoken Utterances. **Link**. *Interspeech 2022.*

6. Yadav, H., **Ghosh, S.**, Yu, Y. & Shah, R. R. End-to-end Named Entity Recognition from English Speech. **Link**. *Interspeech 2020.*

7. **Ghosh, S.**, Seth, A. & Umesh, S. DeLoRes: Decorrelating Latent Spaces for Low-Resource Audio Representation Learning. **Link**. *SAS Workshop at AAAI 2022.*

8. **Ghosh**, **S.** & Kumar, S. Cisco at SemEval-2021 Task 5: What's Toxic?: Leveraging Transformers for Multiple Toxic Span Extraction from Online Comments. **Link**. *SemEval-2021 at ACL 2021.*

9. **Ghosh, S.** *et al.* Cisco at AAAI-CAD21 shared task: Predicting Emphasis in Presentation Slides using Contextualized Embeddings. **Link**. *CAD-21 at AAAI 2021.*

10. **Ghosh, S.**, Seth, A., Umesh, S. & Manocha, D. MAST: Multiscale Audio Spectrogram Transformers. *Under review at ICASSP 2023.*

11. Seth*, A., **Ghosh*, S.**, Umesh, S. & Manocha, D. SLICER: Learning universal audio representations using low-resource self-supervised pre-training. *Under review at ICASSP 2023.*

12. **Ghosh**, **S.** *et al.* M-MELD: A Multilingual Multi-Party Dataset for Emotion Recognition in Conversations. *Under review at ICASSP 2023.*

13. **Ghosh**, **S.**, Seth, A., Katta, S. V. & Umesh, S. Deep Clustering for learning general-purpose Audio Representations. **Link**.

14. **Ghosh**, **S.**, Srivastava, H. & Umesh, S. MMER: Multimodal Multi-task learning for Emotion Recognition in Spoken Utterances. **Link**.

15. **Ghosh**, **S.**, Srivastava, H. & Umesh, S. A Discourse Aware Sequence Learning Approach for Emotion Recognition in Conversations. **Link**.

16. Prasad, L. V. S. V. D., Seth, A., **Ghosh**, **S.** & Umesh, S. Analyzing the factors affecting usefulness of Self-Supervised Pre-trained Representations for Speech Recognition . **Link**.

17. **Ghosh**, **S.**, Sunny, S. J. & Roney, R. Accident Detection Using Convolutional Neural Networks. Link. *IconDSC* (2019).

18. **Ghosh**, **S.**, Sakshi & Kataria, V. Multivariate Time-series Unsupervised Anomaly Detection in 5G networks. Defensive Publication. *Under US Patenting with Cisco* (2020).

Projects (Applied)

Automatic Speech Recognition for Indian Languages   **July 2020 — Present**
- Used state-of-the-art neural networks and self-supervised algorithms to create a multi-lingual ASR system that has sub 10% WER on 7 Indian languages.
- Used a combination of adverserial learning and transfer learning to reach state-of-the-art WER on an ASR system for Indian accented English.
- Implemented several reasearch papers together with developing new algorithms during the course of this project.
- Served ASR systems for real-time and offline speech transcription.
- Part of this work was done at MIDAS@IIIT-D. I am currently continuing building ASR for Indian languages under the National Mission at Speech Lab IIT-Madras.

Intelligent Minutes of the Meeting   **July 2020 — August 2020**
- Used Automatic Speech Recognition and Speaker Diarization together with transformer-based text summarization and Named-Entity-Recognition to create a web app for intelligent Minutes of the Meetings.
- Application has capabilities to take meeting recordings and output summary, diarized transcripts, and key points from the meeting, including dated items and to-dos.

SpeeQL   **June 2020 — July 2020**
- Used Automatic Speech Recognition, NER, and Image Classification together with OCR to build a smart in-store touch-free shopping assistant for brick and mortar stores.
- Project was one of the first to use speech as a medium of interaction. Was later adopted by P&G to deploy on scale.

User fatigue prediction using deep learning   **July 2019 — August 2019**
- Building End-to-End systems using object detection and image classification techniques to detect levels of fatigue in different parts of the face.
- Build a novel dataset for the same. The application was able to give a user suggestions based on his total level of fatigue calculated using weighted average of fatigue from all parts of the face.

Awards & Honors

2020   Recognised by Cisco CX CTO and higher management on multiple occasions for my research and innovation initiatives.
2020   Awarded the **Graham Bell Award** for being one of the most competitive undergraduates to have graduated in the year 2020
2020   Winner of Cisco Collab Hacks
2020   Winner of P&G Global Innovation Challenge
2019   Appeared on the cover page of Analytics India Magazine twice for winning national level hackathons (**TEG Analytics and Uber Hackathon**)
2019   Winner of Hindustan Unilever BFS Technology Hackathon
2018-20 Winner of various inter-college and intra-college hackathons sponsored by MNCs and the Government (Including a bronze medal at Kaggle)

Community Service

2019-20 Served as the Secretary of Computer Science Department, Christ University
2018-20 Co-founded Neuron, Christ University's first AI club focused on research, served as the first Vice President of the club
2022   Reviewer at AAAI 2023, ACL 2021, HASOC 2021, Electronic Letters (IET)
2022   Lecturer of SLP at University of Buffalo, New York.

Co-Curricular and Hobbies

2015   State-Level Table Tennis Champion (West Bengal, India)
Present Active investor and trader in the stock markets. In-depth knowledge of fundamental and technical analysis of stocks.
Present Nature Photography