ARINDAM JATI

✉ arindamjati@gmail.com http://arindamjati.com/

Research Interests

Machine learning, Deep neural networks, Explainable AI, Timeseries forecasting

Education

University of Southern California (USC), Los Angeles, CA, USA 2015 - 2021
PhD in *Electrical Engineering*
Thesis: *Understanding sources of variability in learning robust deep audio representations* [link]
GPA: 3.91/4.0

University of Southern California (USC), Los Angeles, CA, USA 2015 - 2017
Master of Science (MS) in *Electrical Engineering*
GPA: 3.91/4.0

Jadavpur University, Kolkata, India 2009 - 2013
Bachelor of Engineering (BE) in *Electronics and Telecommunication Engineering*
GPA: 9.43/10.0

Work Experience

Research Scientist Apr 2021 - present
IBM Research, Bangalore, India

Graduate Research Assistant Aug 2015 - Mar 2021
SAIL Lab. at University of Southern California (USC), Los Angeles, CA, USA
Advisor: Prof. Shrikanth Narayanan *Past Advisor:* Prof. Panayiotis Georgiou

- ***Self-supervised*** deep speaker representation learning/pre-training
- ***Adversarial attack*** on speaker recognition system, and defense strategies (DARPA GARD)
- Workplace ***acoustic scene*** identification from egocentric data (IARPA MOSAIC)
- Effect of workplace ambience sounds on ***employee behavior*** and ***performance*** (IARPA MOSAIC)
- ***Multi-task*** training of robust speaker embedding
- ***Multimodal*** depression detection from audiovisual data

Research Intern May - July, 2019
Audio and Acoustics Research Group at Microsoft Research, Redmond, WA, USA
Manager: Dr. Ivan Tashev, *Mentor:* Dr. Dimitra Emmanouilidou

- ***Deep semantic hashing*** method for efficient audio event retrieval

AI Intern June - Aug, 2018
Sony Interactive Entertainment, San Mateo, CA, USA
Manager: Dr. Ruxin Chen, *Mentor:* Dr. Naveen Kumar

- Developed a novel framework for learning deep ***audio event*** embeddings for hierarchical label-space
- Built an on-demand ***accessibility*** system that provides sound/action descriptions to users (with hearing/visual impairments) playing videogames

Graduate Teaching Assistant 2017 - 2019
University of Southern California (USC), Los Angeles, CA, USA

- TA experience in graduate-level courses: Deep Learning, Deep Learning Lab for Speech Processing, Mathematical Pattern Recognition, and Digital Signal Processing

Software Engineer - II & I 2013 - 2015
Polaris Networks, Kolkata, India

- Developed node emulators and test tools for 4G LTE communication networks

Undergraduate Research Intern Dec, 2011 to Jan, 2012
School of Medical Science and Technology, IIT Kharagpur, India

- Worked on medical image segmentation using fuzzy sets

Previous Research Experience (Topics)

Machine Learning, Deep Learning, Adversarial Robustness: Deep Representation Learning, Unsupervised & Self-supervised Learning, Hierarchical Representations, Quantized Representations, Multi-task Learning, Adversarial Attack on Deep Neural Nets & Defense Strategies

Audio, Speech & Natural Language Processing: Speech Recognition, Speaker Recognition, Speaker Diarization, Audio Event & Acoustic Scene Identification

AI & Machine Learning for Affective Computing: Multi-modal Human Emotion & Behavior Recognition, Stress & Anxiety Detection, Depression Prevention, Learning from Egocentric Data

Honors and Achievements

- IBM Research India 2021 **Best Repository Award**.
- Thesis work on adversarial attack and defense for speaker recognition systems was mentioned in the **news**.
- Honorable mention for **Best Teaching Assistant** (TA) award, 2019 at USC.
- Honorable mention **(3^{rd} place) in Summer 2018 Hackathon** at Sony Interactive Entertainment America LLC.
- Received **ISCA travel grant award** for students and young scientists for Interspeech 2017 conference.
- Received **Annenberg PhD Fellowship** at USC.

Selected Publications (Full List in Google Scholar)

1. Monisankha Pal, **Arindam Jati**, Raghuveer Peri, Chin-Cheng Hsu, Wael AbdAlmageed, Shrikanth Narayanan, *"Adversarial defense for deep speaker recognition using hybrid adversarial training"*, Accepted in ICASSP 2021. [arXiv]
2. **Arindam Jati**, Chin-Cheng Hsu, Monisankha Pal, Raghuveer Peri, Wael AbdAlmageed, Shrikanth Narayanan, *"Adversarial Attack and Defense Strategies for Deep Speaker Recognition Systems"*, in Elsevier Computer Speech and Language, Volume 68, 101199, 2021. [arXiv] [doi] [pdf]
3. **Arindam Jati**, Amrutha Nadarajan, Raghuveer Peri, Karel Mundnich, Tiantian Feng, Benjamin Girault, and Shrikanth Narayanan, *"Temporal Dynamics of Workplace Acoustic Scenes: Egocentric Analysis and Prediction"*, in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 756-769, 2021. [doi] [pdf]
4. **Arindam Jati**, and Dimitra Emmanouilidou, *"Supervised Deep Hashing for Efficient Audio Event Retrieval"*, In ICASSP 2020. [pdf]
5. Raghuveer Peri, Haoqi Li, Krishna Somandepalli, **Arindam Jati**, and Shrikanth Narayanan, *"An empirical analysis of information encoded in disentangled neural speaker representation"*, in Odyssey: The Speaker and Language Recognition Workshop, 2020. [pdf]
6. Raghuveer Peri, Monisankha Pal, **Arindam Jati**, Krishna Somandepalli, and Shrikanth Narayanan, *"Robust speaker recognition using unsupervised adversarial invariance"*, In ICASSP 2020. [pdf]
7. **Arindam Jati**, Raghuveer Peri, Monisankha Pal, Tae Jin Park, Naveen Kumar, Ruchir Travadi, Panayiotis Georgiou, and Shrikanth Narayanan, *"Multi-task Discriminative Training of Hybrid DNN-TVM Model for Speaker Verification with Noisy and Far-Field Speech"*, In Interspeech 2019. [pdf]
8. Krishna Somandepalli, Naveen Kumar, **Arindam Jati**, Panayiotis Georgiou and Shrikanth Narayanan, *"Multiview Shared Subspace Learning across Speakers and Speech Commands"*, In Interspeech 2019. [pdf]
9. **Arindam Jati**, Naveen Kumar, Ruxin Chen, and Panayiotis Georgiou, *"Hierarchy-Aware Loss Function on a Tree Structured Label Space for Audio Event Detection"*, In ICASSP 2019. [pdf]

10. **Arindam Jati** and Panayiotis Georgiou, *"An unsupervised neural prediction framework for learning speaker embeddings using recurrent neural networks"*, In Interspeech, 2018. [pdf]
11. **Arindam Jati** and Panayiotis Georgiou, *"Neural Predictive Coding using Convolutional Neural Networks towards Unsupervised Learning of Speaker Characteristics"*, in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 10, pp. 1577-1589, Oct. 2019. doi: 10.1109/TASLP.2019.2921890, 2018. [arXiv] [doi]
12. **Arindam Jati**, Paula G. Williams, Brian Baucom and Panayiotis Georgiou, *"Towards Predicting Physiology from Speech During Stressful Conversations: Heart Rate and Respiratory Sinus Arrhythmia"*, In ICASSP, 2018. [pdf]
13. **Arindam Jati** and Panayiotis Georgiou, *"Speaker2Vec: Unsupervised Learning and Adaptation of a Speaker Manifold using Deep Neural Networks with an Evaluation on Speaker Segmentation"*, Proceedings of Interspeech, 2017. [pdf]
14. Md Nasir, **Arindam Jati**, Prashanth Gurunath Shivakumar, Sandeep Nallan Chakravarthula, and Panayiotis Georgiou, *"Multimodal and Multiresolution Depression Detection from Speech and Facial Landmark Features"*, Proceedings of the 6th ACM International Workshop on Audio/Visual Emotion Challenge (AVEC). ACM, 2016. [pdf]

Patents

Granted

1. Ashish Singh, Justice Adams, **Arindam Jati**, Masanori Omote, *"Color accommodation for on-demand accessibility"*, US Patent, 2020. [US20200135052A1]
2. Sudha Krishnamurthy, Ashish Singh, Naveen Kumar, Justice Adams, **Arindam Jati**, Masanori Omote, *"Graphical style modification for video games using machine learning"*, US Patent, 2021. [US20200134929A1]

Filed

1. **Arindam Jati**, Naveen Kumar, Ruxin Chen,*"Sound Categorization System"*, US Patent filed, 2018. [US20200104319A1]
2. Justice Adams, **Arindam Jati**, Sudha Krishnamurthy, Masanori Omote, Jian Zheng, Naveen Kumar, Min-Heng Chen, Ashish Singh, *"Action description for on-demand accessibility"*, US Patent filed, 2018. [US20200129860A1]
3. Sudha Krishnamurthy, Justice Adams, **Arindam Jati**, Masanori Omote, Jian Zheng, "*Scene annotation using machine learning"*, US Patent filed, 2018. [US20200134316A1]
4. Naveen Kumar, Justice Adams, **Arindam Jati**, Masanori Omote, *"Textual annotation of acoustic effects"*, US Patent filed, 2018. [US20200137463A1]

Talks

1. *"Supervised Deep Hashing for Efficient Audio Retrieval"*, at Microsoft Research, Redmond, WA, USA. [Microsoft Research Page][YouTube]

Open Source Software

1. Adversarial attack and defense strategies for deep speaker recognition systems: `https://github.com/usc-sail/gard-adversarial-speaker-id`

Technical Skills

Programming: Python, Bash, C/C++, MATLAB
Machine learning tools: Pytorch, Keras, Tensorflow, Scikit-learn
Big data & distributed computing: PySpark, RAY
Machine learning on clusters: Amazon AWS, Microsoft Azure, USC HPCC
Speech and NLP tools: KALDI Speech Recognition Toolkit, OpenSMILE, OpenFST
OS: Unix, Windows **Other tools:** Docker, Git, LaTeX, SPSS

Professional Activities

Reviewer

- **Journals: 1.** IEEE/ACM Transactions on Audio, Speech, and Language Processing, **2.** IEEE Signal Processing Letters, **3.** IEEE Access, **4.** EURASIP Journal on Audio, Speech, and Music Processing, **5.** Springer Journal of Signal, Image and Video Processing
- **Conferences: 1.** 20th ACM ICMI 2018, **2.** IEEE ICASSP 2021, **3.** AAAI 2021.

Relevant Graduate Courses

Digital signal Processing	Pattern recognition	Algorithms
Probability	Machine learning	Affective computing
Random processes	Natural language processing	Wavelets and graph signal processing

Selected Course Projects

- Wavelets and graph signal processing: *Sparse Representation of Deep Neural Network Embeddings for Speaker Identification* [pdf]
- Affective Computing: *End-To-End Speech Negotiations with Affective Speech Rollout* [pdf]
- Pattern Recognition: *Predicting Readmission of Diabetic Patients from Medical Records* [pdf]
- Machine Learning: *Santander Customer Satisfaction Classification* [pdf]
- Natural language processing: *Automatic Solver for Mad Gab - A Language Game* [pdf]