

# ARINDAM JATI

✉ jati@usc.edu, arindamjati@gmail.com 🌐 <http://arindamjati.com/>

RESEARCH INTERESTS	Machine Learning, Deep Neural Networks, Audio, Speech & Natural Language Processing, Video & Multi-modal Signal Processing, AI for Social Good, Adversarial Attack & Defense
EDUCATION	<div><div><b>University of Southern California (USC), Los Angeles, CA, USA</b> 2015 - present <b>PhD candidate</b> in <i>Department of Electrical and Computer Engineering</i> <i>Current GPA: 3.91/4.0</i></div><div><b>University of Southern California (USC), Los Angeles, CA, USA</b> 2015 - 2017 <b>Master of Science (MS)</b> in <i>Electrical Engineering</i> <i>GPA: 3.91/4.0</i></div><div><b>Jadavpur University, Kolkata, India</b> 2009 - 2013 <b>Bachelor of Engineering (BE)</b> in <i>Electronics and Telecommunication Engineering</i> <i>GPA: 9.43/10.0</i></div></div>
WORK EXPERIENCE	<div><div><b>Graduate Research Assistant</b> Aug 2015 - present <i>SAIL Lab.</i> at <b>University of Southern California (USC), Los Angeles, CA, USA</b> <i>Advisor:</i> Prof. Shrikanth Narayanan <i>Past Advisor:</i> Prof. Panayiotis Georgiou<ul style="list-style-type: none"><li>Self-supervised deep speaker representation learning/pre-training</li><li>Adversarial attack on speaker recognition system, and defense strategies (<b>DARPA GARD Project</b>)</li><li>Workplace acoustic scene identification from egocentric data (<b>IARPA MOSAIC Project</b>)</li><li>Multi-task training of robust speaker embedding</li><li>Multimodal depression detection from audiovisual data</li></ul></div><div><b>Research Intern</b> May - July, 2019 <i>Audio and Acoustics Research Group</i> at <b>Microsoft Research, Redmond, WA, USA</b> <i>Manager:</i> Dr. Ivan Tashev, <i>Mentor:</i> Dr. Dimitra Emmanouilidou<ul style="list-style-type: none"><li>Deep semantic hashing method for efficient audio event retrieval</li></ul></div><div><b>AI Intern</b> June - Aug, 2018 <b>Sony Interactive Entertainment, San Mateo, CA, USA</b> <i>Manager:</i> Dr. Ruxin Chen, <i>Mentor:</i> Dr. Naveen Kumar<ul style="list-style-type: none"><li>Developed a novel framework for learning deep audio event embeddings for hierarchical label-space</li><li>Built an on-demand accessibility system that provides sound/action descriptions to users (with hearing/visual impairments) playing videogames</li></ul></div><div><b>Graduate Teaching Assistant</b> 2017 - 2019 <b>University of Southern California (USC), Los Angeles, CA, USA</b><ul style="list-style-type: none"><li>TA experience in graduate-level courses: Deep Learning, Deep Learning Lab for Speech Processing, Mathematical Pattern Recognition, and Digital Signal Processing</li></ul></div><div><b>Software Engineer - II &amp; I</b> 2013 - 2015 <b>Polaris Networks, Kolkata, India</b><ul style="list-style-type: none"><li>Developed node emulators and test tools for 4G LTE communication networks</li></ul></div><div><b>Undergraduate Research Intern</b> Dec, 2011 to Jan, 2012 <b>School of Medical Science and Technology, IIT Kharagpur, India</b><ul style="list-style-type: none"><li>Worked on medical image segmentation using fuzzy sets</li></ul></div></div>

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, Ambience Sensing

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

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*”, Submitted 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*”, Under review in Elsevier Computer Speech and Language. [\[arXiv\]](#)
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*”, Accepted in IEEE/ACM Transactions on Audio, Speech, and Language Processing.
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. [\[arXiv\]](#) [\[pdf\]](#)
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. [\[US20200135052A1\]](#)

### 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. [\[US20200137463A1\]](#)
5. Sudha Krishnamurthy, Ashish Singh, Naveen Kumar, Justice Adams, **Arindam Jati**, Masanori Omote, “*Graphical style modification for video games using machine learning*”, US Patent filed, 2018. [\[US20200134929A1\]](#)

## 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>

## MAJOR AWARDS

- Honorable mention for **Best Teaching Assistant** (TA) award, 2019 at USC.
- Honorable mention (3<sup>rd</sup> 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.

## TECHNICAL SKILLS

**Programming:** Python, Bash, C/C++, MATLAB  
**Machine learning tools:** Pytorch, Keras, Tensorflow, Scikit-learn  
**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

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* [\[pdf\]](#)