PHAYE SAISAMARTH RAJESH

Audio Data Scientist
Teredo Analytics
LinkedIn: **https://www.linkedin.com/in/ssrp**
Google Scholar: **Ds047PsAAAAJ**
GitHub: **https://github.com/ssrp**

Nationality: **Indian**
DOB: **06 October 1996**
Contact No.: **+65 (0) 9869 3516**
email-id: **phaye.samarth@gmail.com**
Personal: **ssrp.github.io**

WORK EXPERIENCE

- **Senior Audio AI/ML Engineer, Logitech Europe Ltd**
  (Manager: Andrew Harper, since Jan 2024)
  - Developing speech enhancement algorithms for the Logitech's video conferencing devices.

- **Audio Data Scientist, Teredo Analytics Pte Ltd**
  (Founders: Dr. Rajat Mishra and Wong Liang Jie, July 23 - Dec 2023)
  - Developed acoustic systems for numerous predictive maintenance applications using AI/ML.
  - Researched on anomaly detection systems for acoustic based machine faults detection and pipeline leakage localisation.

- **Audio AI Engineer/Researcher, Zoom Singapore**
  (Managers: Cheng-Lun Hu and Alex Liu, Oct 21 - March 23)
  - Worked on the **first AI powered Acoustic Echo Cancellation (AEC) system** for the Zoom platform. The AI system is successfully deployed in Zoom Rooms for the DTEN meetings.
  - Proposed novel neural networks and methods, suitable to **handle numerous real-time meeting scenarios**, such as double-talk, farend-singletalk in long-delay high-reverb echo cases.
  - Devised an algorithm to quantify similarity between multiple audio datasets [**US Patent Submitted**]. Also **created multiple audio datasets** to simulate various meeting scenarios to improve the overall performance of Zoom's AEC system.

- **Machine Learning Research Engineer, Shopee Singapore**
  (Managers: Jingchang Zhang and Dr. Zhenwei Zhao, July 20 - Sept 21)
  - Handled two projects and developed numerous NLP models which shape the backbone of Shopee's recommendation system serving eight countries.
  - Using named-entity recognition to **develop item-attribute prediction models** which are further leveraged to improve the recommendation models.

- **Machine Learning Research Engineer, Panasonic Industrial Devices Singapore, Singapore**
  (Managers: Dr. Vasileios Vonikakis and Dr. Aryel Beck, July 19 - June 20)
  - Developed AI-powered systems for human-sensing technologies for B2B solutions. Worked on face analytics using classical computer vision/deep learning – predicting human demographics using face images in the wild; deploying on the edge devices without compromising the performance. Focusing on **feature selection and learning, model optimization and deployment**.
  - Developed an end-to-end system for unsupervised anomaly detection in the CCTV factory videos to detect abnormalities, such as unauthorized access of machines, machine stopping. Introduced a novel ML algorithm for the use-case which executes in real-time [US Patent].

- **Graduate Researcher, Sound and Music Computing Lab, National University of Singapore** [GitHub]
  (Supervisors: Dr. Emmanouil Benetos and A/Prof. Ye Wang, Aug 18 - June 19)
  - Researched on novel deep neural methods for Acoustic Scene Classification and assisted in related projects. Proposed a novel architecture for acoustic data, **published in IEEE ICASSP 2019**.
  - Teaching Assistant for the course **CS4347: Sound and Music Computing** at School of Computing, National University of Singapore (II Semester AY 2018-19).

ACHIEVEMENTS

- Secured a rank under **top 0.001 percentile** (out of **1.3 million candidates)** in Joint Entrance Exam (Mains and Advanced combined) – national level university entrance examination in India, 2014.
- Won "**Most Interesting Use of AI**" title at the **1st Sound of AI Hackathon**, July 22.
  - Developed **Not Final Cut Pro**, an AI based video mashup application which creates a mashup video using your favorite videos and a song.
- Won "**Hack In Punjab**" - **All India Level Hackathon Event**, organised by **Top Careers And You**, Sept 15.
  - Created an Android Application using **Ionic Framework** which would solve problems faced by travellers. The app alarms the user when (s)he is within 2km radius of the destination.

ENTREPRENEURIAL EXPERIENCE

- **Co-founder, Echolair**
  (Co-founders: Jay Anand and Mahshid Alinoori, April 23 - Dec 2023)
  - Echolair transforms music samples into countless unique variations, empowering music creators to break the constraints of their sound libraries and discover endless inspiration.
  - Interviewed 40+ customers to find the common problem and after multiple pivots, Echolair was born.
  - Developed a fast AI model for variation generator model and the supporting backend to execute the model on the customers' system/laptop.
  - A single 10 second sample takes 0.35s to generate one variation on a common laptop without a GPU. The AI model generates multiple variations of the sample using some tweaks in the latent space.

TECHNICAL SKILLS

- **ML/DL Libraries** - PyTorch, Keras, TensorFlow, scikit-learn, OpenCV, ONNX
- **Programming/Scripting Languages** - Python, Golang, JavaScript, JAVA, C, C++, SQL
- **Tools** - Hadoop, Jupyter Notebook, MATLAB, LaTeX

PUBLICATIONS

- **[Journal Article] José Ignacio Orlando, .., Sai Samarth R Phaye, .., Hrvoje Bogunović. REFUGE Challenge: A Unified Framework for Evaluating Automated Methods for Glaucoma Assessment from Fundus Photographs. Accepted in *Medical Image Analysis, Volume 59, January 2020*** [Report] [Paper]
  - Proposed a 2-level model in the challenge – first sub-network localizes the optic disc region in the sample fundus image, which is further used by second sub-network for finer segmentation of cup/disc.
  - Being **one of the top ten teams**, got invited to participate in the onsite round and orally present our work in Medical Image Computing and Computer Assisted Intervention (MICCAI) 2018, Spain.
- **[Conference] Sai Samarth R Phaye, Emmanouil Benetos and Wang Ye. SubSpectralNet – Using Sub-Spectrogram based Convolutional Neural Networks for Acoustic Scene Classification. Accepted in *14^{th} IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2019*** [arXiv] [GitHub] [PPT]
  - Proposed SubSpectralNets – a deep learning architecture for acoustic scene classification, which leverages band-wise temporal information of the input time-frequency representations.
  - SubSpectralNet achieves a relative **improvement of +14% accuracy** over the baseline model by Detection and Classification of Acoustic Scenes and Events (DCASE) 2018.
- **[Conference and Workshop] Sai Samarth R Phaye*, Apoorva Sikka*, Abhinav Dhall and Deepti Bathula. Multi-level Dense Capsule Networks. Accepted in *14^{th} Asian Conference on Computer Vision (ACCV2018), Perth, Australia*, 2 - 6 December 2018. Also accepted in *Women in Machine Learning (WiML) Workshop, NeurIPS 2019, Vancouver, Canada*, 9 December 2019.** [arXiv] [GitHub]
  - Proposed Multi-level DCNets – introduces multiple level of capsules in Capsule Networks (CapsNets, NIPS 2017) and empowers the complete network using DenseNets, CVPR 2017.
  - Achieves **state-of-the-art performance on MNIST dataset** with **20-fold reduction of training iterations**; **7-fold decrease in the number of parameters** on the CIFAR-10 dataset over the CapsNets.
- **[Technical Demo] Sai Samarth R Phaye, Love Mehta and Mukesh Saini. The One Man Show. Accepted in *19^{th} IEEE International Symposium on Multimedia (ISM2017), Taiwan*, 11 - 13 December 2017** [Paper]
  - Developed an android application helpful for solo-musicians. Lets you record various single instrument videos in sync and using the proposed algorithm, merge them intelligently to create a final video.
  - Uses a **multi-modal system** (SVMs for audio and MLP for image) along with multimedia processing for **instrument recognition in the videos** (for instrument-based audio enhancements such as reverb).

EDUCATION

Qualification	University	Institute	Year	Grade/%
B. Tech.	IIT Ropar	IIT Ropar, India	2014-2018	7.28/10
Intermediate/+2	CBSE	HHEA Raipur, India	2013-2014	91.6/100
Matriculation	CBSE	HHEA Raipur, India	2011-2012	9.6/10

RELEVANT COURSEWORK

- Audio Content Analysis, Computer Vision, Multimedia Systems, Database Systems, Digital Image Processing, Data Structures and Algorithms, Discrete Mathematics, Linear Algebra, Probability Theory, Computer Graphics, Computer Architecture, Programming Paradigms and Pragmatics, Operating Systems

OTHERS

- **When not working**, I am **composing music** and here is my **YouTube Channel**. Ex-member of Guitar Ensemble at National University of Singapore (**GENUS**).

REFERENCES

- **Wong Liang Jie**, Co-founder, Teredo Analytics Pte Ltd, Singapore
    - Contact: +65 (0) 9797 0873, wlj@teredo.sg
- **Cheng-Lun Hu**, Team Lead and Manager, Audio AI Engineering, Zoom Video Communications, Singapore
    - Contact: +65 (0) 8712 2058, chenglun.hu@zoom.us
- **Dr. Emmanouil Benetos**, Senior Lecturer, School of EECS of Queen Mary University of London, UK
    - Contact: +44 (0) 20 7882 6206, emmanouil.benetos@qmul.ac.uk
- **Dr. Deepti Bathula**, Assistant Professor, Indian Institute of Technology Ropar, India
    - Contact: +91 (0) 95011 96606, bathula@iitrpr.ac.in
- **Dr. Abhinav Dhall**, Lecturer, Human-Centred Artificial Intelligence lab, Monash University, Australia
    - Contact: +61 (0) 403 340 928, abhinav.dhall@monash.edu
- **Dr. Vasileios Vonikakis**, Senior Principal Engineer, Panasonic Industrial Devices Singapore
    - Contact: +65 (0) 8649 3499