

# Jayneel Parekh

Postdoctoral Researcher at Sorbonne Université

Research areas: ML/AI, Interpretable ML, Computer vision, Audio processing

✉ [parekh@isir.upmc.fr](mailto:parekh@isir.upmc.fr)

📁 [jayneelparekh.github.io](https://github.com/jayneelparekh)

🌐 [linkedin.com/in/jayneel-parekh/](https://www.linkedin.com/in/jayneel-parekh/)

## Current Position

December 2023–Present **Postdoctoral Researcher**, SORBONNE UNIVERSITÉ France  
[Understanding Representations in Large Multimodal Models](#)

**Advisor:** Prof. Matthieu Cord, (Sorbonne Université and Scientific director of valeo.ai)

- Developing methods to understand representations computed inside large vision-language models.
- Leveraging insights gained from our interpretability methods to control or steer model outputs.

## Education

September 2019–July 2023 **PhD in Machine Learning**, TÉLÉCOM PARIS, Institut Polytechnique de Paris, France

**Thesis:** A Flexible Framework for Interpretable Machine Learning: Application to Image and Audio Classification

**Advisors:** Prof. Florence d'Alché-Buc & Prof. Pavlo Mozharovskyi, Telecom Paris

**Jury:** Grégoire Montavon, Patrick Pérez, David Alvarez-Melis, Nicolas Thome, Stéphane Canu, Chloe Clavel

- Developed post-hoc/by-design interpretation methods using concept-based representations for deep neural networks, with applications to audio and image classification tasks

July 2014–19 **Dual Degree (B.Tech + M.Tech)**, IIT Bombay, Electrical Engineering, CPI: 9.05/10

**Thesis:** Audio Style Transfer: Transformations between speech and singing

**Advisors:** Prof. Preeti Rao, IIT Bombay & Dr. Yi-Hsuan Yang, Academia Sinica, Taiwan

- Investigated style transfer techniques for audio signals
- Focused on converting spoken audio into sung audio and vice-versa using deep learning

## Awards and Achievements

2023 **Top reviewer (top 10%)** for NeurIPS'23.

2023 **STIC Doctoral Prize (Saclay) 2023 accessit (2nd place)** for NeurIPS'22 publication.

2016 **Distinctive mention and travel grant** for work at MediaEval 2016

2013 **Kishore Vaigyanik Protsahan Yojana (KVPY) Fellowship 2013**

## Selected Publications/Preprints

Complete list of papers on [Google Scholar](#). \* denotes equal contribution below

- [1] P. Khayatan\*, M. Shukor\*, **J. Parekh\***, and M. Cord. "Analyzing Fine-tuning Representation Shift for Multimodal LLMs Steering." **arXiv 2025** (under submission). [\[paper\]](#)
- [2] G. Kasmi, A. Brunetto, T. Fel and **J. Parekh**. "One Wave To Explain Them All: A Unifying Perspective On Feature Attribution" **ICML 2025** (Accepted). [\[paper\]](#)
- [3] **J. Parekh\***, Q. Bouniot\*, P. Mozharovskyi, A. Newson and F. d'Alché-Buc. "Restyling Unsupervised Concept Based Interpretable Networks with Generative Models." **ICLR 2025**. [\[paper\]](#)
- [4] **J. Parekh**, P. Khayatan, M. Shukor, A. Newson and M. Cord. "A Concept-Based Explainability Framework for Large Multimodal Models." **NeurIPS 2024**. [\[paper\]](#)
- [5] **J. Parekh**, S. Parekh, P. Mozharovskyi, G. Richard and F. d'Alché-Buc. "Tackling Interpretability in Audio Classification Networks with Non-negative Matrix Factorization." **IEEE/ACM TASLP** (Vol. 32), arXiv 2023. [\[paper\]](#)
- [6] **J. Parekh**, S. Parekh, P. Mozharovskyi, F. d'Alché-Buc and G. Richard. "Listen to Interpret: Post-hoc Interpretability for Audio Networks with NMF." **NeurIPS 2022**. [\[paper\]](#)
- [7] **J. Parekh**, P. Mozharovskyi & F. d'Alché-Buc. "A Framework to Learn with Interpretation." **NeurIPS 2021**. [\[paper\]](#)
- [8] **J. Parekh**, P. Rao, and YH. Yang. "Speech-to-Singing Conversion in an Encoder-Decoder Framework." **IEEE ICASSP 2020 (Oral)**. [\[paper\]](#)

---

## Undergraduate Research Experience

For my graduate or later research experience, please see the **Education** section

- Feb–April 2018 **Supervised Research Exposition**, IIT Bombay  
**Surface Defect Detection**  
**Advisor:** Prof. Subhasis Chaudhuri, Department of Electrical Engineering, IIT Bombay
- Explored various techniques and studied relevant literature for surface defect detection
  - Applied transfer-learning based methods for plant disease detection
- May–June 2017 **Summer Intern**, TECHNICAL R&D France  
**2D & 3D Human Pose Estimation Networks**  
**Advisors:** Pierre Hellier (Principal Scientist) & Louis Chevallier, Technicolor R&D France
- Completed internship as part of an industrial project on Motion Synthesis in Animation
  - Implemented and tested stacked hourglass based deep CNNs for 2D and 3D human pose estimation
- 2016 & 2017 **MediaEval Benchmarking Initiative**  
**Predicting Media Interestingness Task** (**Organizers:** Technicolor France, ETH Zurich *et al.*)
- Developed novel methods for ranking a set of images/video-shots extracted from movie trailers according to their interestingness to a common viewer
  - Netherlands 2016: achieved a mean average precision (MAP) of 0.23 for images - **Team Rank 3/12**
  - Ireland 2017 (web participation): achieved a MAP of 0.25 for images & 0.19 for videos
- 2016–18 **Selected Course Projects**
- **Image Style Transfer** – An unsupervised, graph signal processing based, random shallow CNN.
  - **Blind Audio Source Separation** – Implemented a NMF, LPC based error clustering criterion algorithm
  - **Detection of Moving Objects in Videos** – Implemented a mean-shift and max-flow min-cut based algorithm

---

## Services

- Organizing** ICCV 2025 workshop on explainable Computer Vision (exCV), Trustworthy and Frugal ML workshop, ELLIS Unconference 2023 Paris (Co-chair), ELLIS Member,
- Reviewing** ICCV'25, CVPR'25, ICLR'25, NeurIPS'24, ICML'24, NeurIPS'23, TPAMI, JMLR, TMM
- Courses** Teaching assistant for: **Machine Learning** (Telecom Paris, thrice during 2020–22), **Analytical Signal Processing** (IIT Bombay, Spring 18-19), **Probability & Random Processes** (IIT Bombay, Autumn 18-19) and **Linear Algebra** (IIT Bombay, Autumn 17-18)

---

## Skills

- Programming** Languages: Python (including PyTorch, Tensorflow, Keras), C#, C++
- Courses** Advanced courses in ML/AI, CV, Optimization, Signal Processing, Summer schools – MuSTeR 2016 (IISc Bangalore), MLSS 2021 (Virtual), OxML 2022 (University of Oxford)

---

## Selected Talks

- Jan 2025 IIT Jodhpur – “The Model Interpretability Files: Uncovering Hidden Secrets with Dictionary Learning”
- Apr 2024 Invited talk at ICASSP'24 workshop on XAI for Speech and Audio – “Tackling Interpretability Problems for Audio Classification Networks with NMF”
- Sept 2023 Deezer Research Paris – “Listen to Interpret: Post-hoc interpretability for Audio Networks with NMF”
- 2021–2022 Talks at IDEMIA and Datacraft Paris – “A Framework to Learn with Interpretation”

---

## Extra-Curricular

- 2014–18 **Silver medal** in Institute Squash League, **3X finalist** of Squash General Championship
- 2007–09 **Participated** in Junior National Squash Championship, Otters Open, CCI Open, HC Golcha Memorial Rajasthan Open Squash Championship, Umed Club District Open Squash Championship (**Finalist**)
- 2008–12 Member of Choir group in DPS Jodhpur
- 2009 & 2005 Samvaad – A personality development program – 45 and 30 days respectively