

# Jay Sawant

☎: +1 2068507888 ✉: jsawant@ucsd.edu 🔗: linkedin.com/in/jaysawant61 🐙: jay6101.github.io 🌐: jay6101 🎓: jay-sawant

## Education

### University of California San Diego

San Diego, CA

- MS in Data Science student at Halicioğlu Data Science Institute (HDSI) | CGPA - 4/4 [Expected Sep '24 - Dec '25]
- Key courses: Reinforcement Learning, Deep Generative Models, Machine Learning, Statistical Models, NLP in Biomedicine

### Indian Institute of Technology Bombay

Mumbai, India

- Graduated with a Bachelor's and Master's Degree in Electrical Engineering | CGPA - 8.9/10 [Jul '18 - Jun '23]
- Relevant courses: Machine Learning, Generative AI with LLMs (Coursera), Automatic Speech Recognition, Probability and Random Process, Linear Algebra, Medical Image Computing, Advanced Image Processing, Data Structures and Algorithms

## Technical Skills and Extracurricular

**Programming & Tools** Python, PyTorch, CUDA, Tensorflow, AWS, GCP (Vertex AI), HuggingFace, Pandas, Git, SQL

**Teaching Assistantship** Introduction to Machine Learning (EE769), Introduction to Digital Image processing (EE610)

## Publications

- Patil, A.; Diwakar, H.; **Sawant, J.**; Kurian, N.C.; Yadav, S.; Rane, S.; Bameta, T.; Sethi, A. Efficient Quality Control of Whole Slide Pathology Images with Human in-the-Loop Training. *J. Pathol. Inform.* **2023**, *14*, 100306

## Work, Research and Internship Experience

- **3D Brain MRI Synthesis using DiT-3D** | *Graduate Student Researcher* | *UC San Diego Health* [Jan '25 - Present]
  - Developed a novel conditional latent diffusion pipeline by integrating a VAE for latent space compression with a DiT-3D based diffusion transformer, enabling the generation of high-fidelity, class-specific synthetic 3D brain MRIs with FID score of 16.9
  - Augmented the real brain MRI dataset with synthetic data, leading to substantial improvements in F1-score, accuracy, AUC-ROC, and AUCPR for a Temporal Lobe Epilepsy classifier trained using an EfficientNet-V2 backbone.
- **Cardiac arrhythmia detection and classification** | *Probeplus Innovative Solutions Pvt. Ltd.* | *AI Consultant* [May '24 - Jul '24]
  - Trained a baseline CNN and multi-head attention based model as a part of a remote-health monitoring system to achieve a multi-label classification of 26 arrhythmias using a diverse dataset of 12-lead ECG recordings from Physionet 2021 challenge
  - Improved the model performance by integrating an RNN branch for temporal context and by calculating cross-attention features between CNN and RNN outputs, achieving 2% increase in the challenge score on external test datasets
- **Software Development Engineer** | *Enphase Energy, Bangalore, India* [Jul '23 - May '24]
  - Collaborated with a 7-member Test Automation team to develop and maintain a Python-based test framework
  - Utilized Object-Oriented Programming (OOP) techniques to create comprehensive test suites for the hardware test automation
  - Developed a Python library leveraging the Jama REST API to connect to Jama, fetch test cases, execute them on a local PC, and update results in Jama, providing end-to-end automation support for the Design-Verification-Test team
- **Improving Histopathology and Medical Image Analysis with Deep Learning** | *M.Tech Thesis* [May '22 - Jun '23]
  - Employed active learning method to train a classifier, achieving efficient segmentation of WSIs into six tissue regions and outperforming the popular HistoQC tool with higher dice scores on 70% of the WSIs
  - Utilized DeepLabV3 architecture for cell detection and classification using segmentation method in the OCELOT Challenge 2023, securing a global ranking of 16th place with an F1-score of 0.67
- **Opacity Detection in Chest Xrays using Contrastive Learning** | *Qure.ai* | *AI Scientist Intern* [May '22 - Aug '22]
  - Trained a vanilla classification model of ResNet50 on 1.2M chest x-rays using the conventional supervised training method for opacity classification and achieved an AUC score of 0.80 on the test dataset of 280K chest x-rays
  - Outperformed the vanilla baseline by utilizing a model backbone trained using the Supervised Contrastive Learning method

## Key Technical Projects

- **Cross-Person Virtual Try-On** | *ECE285: Deep Generative Models* | *Guide: Prof. Pengtao Xie* [May'25 - June'25]
  - Leveraged IDM-VITON preprocessing to generate dataset for cross person garment transfer along with DensePose segmentation maps to convert VITON-HD images into cloth-agnostic and garment-conditioned token streams for diffusion-transformer training
  - Architected and implemented a 16-block Diffusion Transformer denoiser integrating self-attention on noise and cloth-agnostic tokens, cross-attention on garment tokens, and FiLM-based timestep modulation for high-fidelity virtual try-on
  - Optimized the diffusion pipeline with noise-aware EDM parameterization, a 2<sup>nd</sup> ODE solver and CFG achieving an FID of 27.7
- **RAG-based Multimodal Medical QnA App** | *MED277: Biomedical NLP* | *Guide: Prof. Shamim Nemati* [Oct'24 - Nov'24]
  - Designed a RAG-based multimodal medical chatbot app with multilingual capabilities, integrating a Google Translate API, BiomedBERT embeddings for context retrieval, Google's Gemini 1.5 Pro for output generation, and Gradio for a web interface
  - Built a pipeline that combines text, medical imaging, and contextual retrieval to deliver accurate responses to medical queries
- **Identity Aware Portrait Generation** | *CS726: Advanced Machine Learning* | *Guide: Prof. Sunita Sarawagi* [Feb'22 - April'22]
  - Utilized the CycleGAN model in the Image translation to generate portraits preserving the human facial features
  - Proposed a perceptual loss to preserve facial features that uses FaceNet embeddings to guide the generators
  - Achieved an average SSIM of 0.98 using our approach between the human faces and their respective portraits