SANKET M. CHAUDHARY

sanketjudal@gmail.com | Linkedin | GitHub

TECHNICAL SKILLS

Languages: Python, R Programming
Software: VS Code, Jupyter Notebook, Google Collab, GitHub, Weka, AutoDock, PLINK, WSL
Deep Learning Framework: Tensorflow, Keras, PyTorch, Transformers, Huggingface, MONAI
Library: OpenCV, OS, NumPy, pandas, Pillow, scikit-learn, Matplotlib, Seaborn, Nibabel, OCR Libraries
Cloud: AWS (Beginner: Lambda, S3, EC2, AWS Bedrock)
AI Services & APIs: OpenAI API, Anthropic Claude, Hugging Face Models, Google Vertex AI

EXPERIENCE

Computer Vision Intern **Apr 2025 - Jun 2025**
Claw Legaltech

- Developed an **autonomous Legal PDF text extracting pipeline** from different regional languages, translating to English, creating a summary, and appending to CSV files, using Huggingface open-source models.
- Created a **multimodal Video summarization pipeline** that **extracts** and **summarizes both** visual (**text** from **video** frames) and **audio**, with a combined summary at the end, using **Whisper to transcribe speech**.
- Worked on the Textile Image generator bot, tried to **fine-tune open-source diffusion models** and AWS Titan G1 model for Image-to-Image and text-to-image generation, **developed Captcha Extractor & Solver** (Audio, Image).

Machine Learning Research Intern **Dec 2024 - Mar 2025**
Design Innovation Centre (DIC) - Hub GTU

- Built a **multimodal Alzheimer's disease classification pipeline** using **ADNI T1-weighted MRI** and **genetic SNP** data for early diagnostic prediction.
- Preprocessed MRI data using **MONAI**, genetic data using **PLINK**, **PCA**, and a **noisy autoencoder** for embedding extraction.
- Fine-tuned **MedicalNet 3D-ResNet50** for MRI feature learning and performed **multimodal fusion with Random Forest**, achieving effective integration of imaging and genetic modalities.

Post-Graduation Dissertation Project **Dec 2023 - Jun 2024**

- Wrote & presented at the international conference 'CDBBI2025', Thesis titled: "**Multimodal Breast Cancer Classification with Medical-Based Transfer Learning**".
- Developed variations of **multimodal CNN algorithms (feature level fusion)** by using **RadImageNet-based pre-trained models** (trained on medical images) and publicly available breast image datasets (MRI, Ultrasound, and Thermal) for breast cancer classification.
- Learned to **preprocess images, fine-tune & hyperparameter-tune the models**, model inference, introduced to CUDA & CentOS to use HPC server, and understood the ethical concerns for medical implications of AI.

Smart India Hackathon (SIH) 2023 Problem Statement: *"AI-based PG Dissertation Management System"*

- As a member of the team 'Rebooters', I contributed to creating the best match profile ranking prediction model using the Random Forest algorithm. I was introduced to Feature Engineering and learned to pitch ideas effectively.

In-Xitu: A student-run entrepreneurial venture **Apr 2022 - May 2023**

- Led a team of four people to manage Tissue culture-based plant production as Head of Production.
- Developed a new plant tissue culture-based product in collaboration with the R&D department.

EDUCATION

Post Graduate Diploma, Bioinformatics (2024 Passed), **Gujarat Technological University**

- **Coursework:** Statistics, Python, Cloud Computing, Data Mining & Visualization, Machine Learning for Life Science, Computer-Aided Drug Design

CERTIFICATES

- Improving Deep Neural Networks: Hyperparameter Tuning, Regularization, and Optimization from DeepLearning.AI.
- AWS Foundational Cloud Computing Bootcamp from AWS Community, Ahmedabad
- Mathematics for Machine Learning Specialization from Imperial College London through Coursera (pursuing)
- ML Ops Fundamentals – Building, Deploying, and Scaling AI Solutions from Scaler