J Vijayavallabh

GitHub | LinkedIn | be23b041@smail.iitm.ac.in | +91 93844 77762

Summary

Pre-final year **B.Tech (Biological Engineering), IIT Madras**; focused on **AI for biotech** and **multimodal/agentic systems**.

Education

2027 **B.Tech Biological Engineering**, **IIT Madras** (CGPA: 8.64/10)
2023 Class XII (CBSE), **Sri Chaitanya Techno School** (95.4%)
2021 Class X (CBSE), **Sri Chaitanya Techno School** (99.4%)

Publications and Patent

- **An Empirical Audit of k-NAF Budget Accounting for Anchored Decoding** (arXiv:2605.28001 preprint).
- **Enhancing Financial RAG with Agentic AI and Multi-HyDE** (Proceedings of **EMNLP** 2025 FinNLP) [Inter-IIT].
- **Sparse Hyperbolic Convolutional Networks with Enhanced Object Localization via GradCAM** (Proceedings of International Conference on Computer Vision **ICCV** Beyond Euclidean Workshop **BEW** 2025) [SENAI IITM].
- **Patent (granted):** SYSTEMS AND METHODS FOR AGENTIC BASED APPROACH IN RETRIEVAL AUGMENTED GENERATION

Professional Experience

AI Genomics Research Intern, Erdős AI Lab Jul 2026
- Developed **EG-IPG** (evidence-gated GNN) over an **8M-edge protein-interaction graph**, integrating **ESM-2** and **PINNACLE** embeddings to predict **CRISPRi Perturb-seq** response in $CD4^+$ T cells (˜11.5K targets) with edge-level evidence gating.
- Built a **training/eval pipeline** (multi-GPU sweeps, test suite, run registry) and ran preregistered multi-seed ablations on leakage-safe target-OOD splits with family-wise error control, yielding a bounded null for an **AAAI 2027** submission.

AI Research and Development Intern, Photios Apr 2025 - May 2025
- Fine-tuned state-of-the-art Automatic Speech Recognition (**ASR**) models using Deep Learning techniques for transcribing sales interactions in retail environments, by extensive literature review and comparative analysis of ASR architectures.
- Architected an automated data cleaning pipeline with the **FastText** model to filter Sanskrit words from Hindi datasets, then fine-tuned custom **N-Gram language model** integrated with **AI4Bharat Conformer**, reducing Word Error Rate (**WER**).

Research Experience

Adversarial Auditing of K-NAF Privacy [Privacy in AI Course Project] Jan - May 2026
- Built an adversarial auditing pipeline for privacy-aware text generation to test K-NAF decoding guarantees under targeted prompt attacks across optimization, validation, stress tests, and held-out evaluation over multiple global K budgets and model variants.
- Measured certified privacy risk via U_{EBB}, effective budget, and ρ; both settings stayed certified in aggregate but $k = 3$ showed stronger held-out transfer, sharper attack separation curves, and more reliable audit validity than $k = 5$ overall.

Research Intern, Kansas State University [Paper under review] May - Jul 2025
- Developed **Position-Aware Inductive Graph Transformer (PAIGT)** for scRNA-seq (Prof. Bala Natarajan), GraphSAGE with degree/shortest-path/path-attention encoders and 6 biological-prior modules (18 CLI controls) over 4 injection points.
- Built leak-free pipeline (split-before-HVG; train-only HVG/kNN/prior std.) with **scTEP** ensemble inference; ran a **12,576-config** × **5-seed** sweep (23 arms, 7 metrics, 88,032) on 4 GPUs across 18 cohorts, boosting acc. **+4.3 pp** and ARI/NMI **+0.59/+0.48**.

Student Researcher, SENAI IITM Dec 2024 - May 2025
- Worked on **geometric representation learning** (Prof. Raghunathan Rengaswamy): Hyperbolic Grad-CAM, activation sparsity in Hyperbolic/hybrid CNNs for interpretability in non-Euclidean spaces, revealing hierarchical/part-whole structure in images.
- Delivered **technical talks**: geometric deep learning, disentanglement via topology, topology of DNNs, and PCA in ML research.

Student Researcher, RBCDSAI Sep 2024 - Dec 2024
- Modelled **reaction-protein relationships** for reaction prediction (Prof. Karthik Raman) by building a **seq2seq** pipeline with **ESM**-encoded proteins and a **BART**-style cross-attention decoder to generate all SMILES sequences for a given reaction sequence.
- Optimized the model capabilities through tokenization and data partitioning experiments to maximize chemically valid SMILES.

Personal Projects

Inter IIT Tech Meet 14.0 (ISRO GeoNLI) Nov - Dec 2025
- Led a 6-member team representing IIT Madras in a **Mid Prep PS** over a span of 15 days to build ***GeoNLI***, an offline/on-prem remote-sensing Vision Language Model [**VLM**] system for **captioning**, semantic/binary/numeric **VQA**, and visual **grounding**.
- Curated **multi-resolution RS** datasets (VRSBench, RSVLM-QA, XLRS) and synthesized adversarial + counting/area VQA.
- Finetuned **Ovis2.5-9B** using **LoRA** on captioning and VQA, and implemented a **hybrid grounding** stack [also in similar fashion for VQA] (Ovis intent routing → **SAM-3** segmentation → OpenCV mask-metadata → OBB + tool-based geometry).

Inter IIT Tech Meet 13.0 (Pathway) Oct - Dec 2024
- Represented IIT Madras in a **High Prep PS** to build a novel dynamic agentic Retrieval Augmented Generation [**RAG**] pipeline for financial documents like **SEC-10K** filings with **adaptive, secure, explainable** workflows using Pathway's vector store.
- Engineered **Multi-HyDE** hybrid retrieval mechanism combining multi-query and hypothetical document embeddings, robust **PDF parsing**, Llama Guard-based **safety guardrails**, and **explainable** AI source attribution; deployed on **Azure** (Docker + FastAPI) and evaluated on **GPT-4o/Gemini-1.5** using semantic similarity, recall, factual correctness, faithfulness, and ROUGE.

Scholastic Achievements

- **Top 1%** in JEE Main and JEE Advanced 2023
- **Reviewer:** UniReps (**NeurIPS** 2025 Workshop)
- **YC Startup School India (Bengaluru)** – Received **YC AI Stack** benefits (*$25k+* in AI tools + cloud credits)
- **FAST Grants (Biswas Family Foundation)** – **$25,000** Fast Grant (Cycle 1; decision *Jul 2026*, funded *˜Aug 2026*)
  PhageHost-Bench: open benchmark for AI-driven phage–host prediction (AI ∩ health; 12-month project)
- **Top 6/23** – Inter IIT Tech Meet 13.0 (High Prep)
  Dynamic Agentic RAG for SEC-10K filings
- **Top 8/23** – Inter IIT Tech Meet 14.0 (Mid Prep)
  ISRO GeoNLI
- **ABRSM Grade 8 Piano Practical** – Merit (Nov 2024)

Skills

Technical	**Python**, Matlab, **PyTorch**, TensorFlow, Pandas, NumPy, RDKit, DGL, NetworkX, **LangChain**, Unsloth, MNE-Python, Optuna, NLTK, Matplotlib, Git/GitHub, Docker, FastAPI, Bash, Azure, Pix2pixHD, LaTeX
Courses/Certs	Protein Structures and Functions, Computational Systems Biology, Bioinformatics, Privacy in AI, Computational Neuroscience, Microbiology and Biochemistry, Principles of Neuroscience, DSA for Biology, Molecular Biology and Genetic Engineering, Biological Engineering Lab, Advanced Topics in AI [Large Scale Cloud Computing and Federated Learning], Signals and Systems, Linear Algebra, Probability and Statistics.