Anshul Bagaria, BE21B005
Indian Institute Of Technology Madras, India
LinkedIn Github Portfolio Email

Education

Indian Institute of Technology Madras — October 2021 – Ongoing
B.Tech. (Honors) in Biological Engineering — **8.85*/10.00** — Chennai, Tamil Nadu
M.Tech. in IDDD Data Science

RV Pre-University College — July 2019 – July 2021
Senior Secondary Class XII — **10.00/10.00** — Bangalore, Karnataka

Research Publications

- **Anshul Bagaria** (2025). "**AMR-MoEGA**: Antimicrobial Resistance Prediction using Mixture of Experts and Genetic Algorithms." (Targeting **ISMB 2026 Proceedings** Submission, to be published in **Bioinformatics**). Available on **arXiv**: **2511.12223**.
- **Anshul Bagaria** (2025). "**INSIGHT**: An Interpretable Neural Vision-Language Framework for Reasoning of Generative Artifacts." (Targeting **ICML 2026** Submission, to be published in **Trustworthy Machine Learning**). Available on **arXiv**: **2511.12223**.

Research & Professional Experience

Quantifying and Mitigating Severity Bias in Medical Large Language Models — **Jul 2025 – Ongoing**
Guide: Dr. Balaraman Ravindran — Masters Thesis, Centre for Responsible AI (CeRAI) — *IIT Madras, Chennai, India*
- Investigating **bias** in clinical LLMs using oncology narratives from **MIMIC-III** to assess fairness and factual fidelity
- Developed an **Oncology Severity** Glossary and graph-based extraction pipeline to map contextual severity patterns.
- Proposed **severity-conditioned attention** and **contrastive representation** learning for severity-aware modeling.
- Incorporating **severity calibration** heads and **reward-guided fine-tuning** for bias mitigation and interpretability.

INSIGHT: Multimodal Artifact-Guided Detection of AI-Generated Images — **Oct 2024 – Dec 2024**
Inter-IIT Tech Meet 13.0, Adobe Research Team AI Challenge — *IIT Bombay, Mumbai, India*
- Designed a framework combining visual and linguistic reasoning, achieving **90% accuracy** on the CIFAKE dataset.
- Integrated **GradCAM**-based artifact localization and **LLM explainability (MOLMO)** for interpretable detection.
- Enhanced robustness through adversarial defense ensemble and knowledge distillation, reducing vulnerability by **21%**.

Multimodal Simulation of User Behavior and KPI-Driven Content Generation — **Oct 2023 – Dec 2023**
Inter-IIT Tech Meet 12.0, Adobe Research MDSR Team AI Challenge — *IIT Madras, Chennai, India*
- Developed a multi-stage **XGBoost pipeline** for robust prediction under cross-**brand** and **temporal** domain shifts.
- Integrated **Mistral-7B** with **LanguageBind embeddings** in **KPI-aware RAG** framework for content generation.
- Built a **vector-indexed KPI database** enabling more efficient semantic prompt retrieval via **cosine similarity**.

Generative AI-Driven Super-Resolution for Lunar Terrain Mapping — **Sept 2022 – Feb 2023**
Inter-IIT Tech Meet 11.0, ISRO Chandrayaan-2 Orbiter Imaging AI Challenge — *IIT Madras, Chennai, India*
- Developed a **super-resolution framework** (SRUN + SORTN) to generate 30 cm-resolution from 10 m TMC data.
- Implemented **spatial attention** U-Nets, and adaptive histogram scaling to ensure high-fidelity image reconstruction.
- Achieved **PSNR 28.26, SSIM 0.79** at 4× upscaling, enabling the generation of preliminary **AI-based lunar atlas**.

Unsupervised Cross-Modality Adaptation for Brain Tumor MRI Segmentation — **Aug 2023 – Aug 2024**
Guide: Dr. Arun K. Thittai — Young Research Fellowship — *IIT Madras, Chennai, India*
- Engineered **intensity-mapping** and **correlation-aware** augmentations to counter **bias** and modality induced shifts.
- Designed **causality-guided** mechanisms to disentangle **spurious correlations** and improve cross-modal transfer.
- Integrated joint **image–feature** adaptation with **nnUNet**, achieving Dice scores of **0.63 (VS)** and **0.60 (Cochlea)**.

Domain Invariant Multi-Organ Segmentation via Contrastive Adaptation — **May 2023 – Dec 2023**
Guide: Dr.Vaanathi Sundaresan — Biomedical Image Analysis(BioMedIA) Laboratory — *IISc, Bengaluru, India*
- Investigated **test-time domain adaptation** frameworks for multi-organ segmentation under cross-scanner shifts.
- Integrated **contrastive alignment** with **transformer encoders** to learn domain-invariant anatomical features.
- Leveraged **adversarial** image translation to enhance invariance and segmentation stability under distribution shifts.
- Achieved steady Dice scores (**0.54–0.80**) across multi-organ CT segmentation benchmarks despite domain disparity.

Multimodal Learning for AI-Driven Diabetic Retinopathy Diagnosis **Jan 2022 – Mar 2023**
ML Developer Intern — SiddhaAI *McKinney, Texas, USA*

- Built a **multimodal diagnostic framework** combining retinal imaging and physiological signals for DR detection.
- Leveraged **semi-supervised SCAN** for efficient label propagation, enhancing generalization on limited clinical data.
- Achieved **mAP50 > 0.98** in vital sign extraction using **YOLOv8** and **OCR-based** monitor text recognition.

Causality Driven Uplift Modeling for Customer Engagement Optimization **May 2025 – Jul 2025**
Consumer Model Development Center — Business Analytics, Wells Fargo *Bengaluru, Karnataka, India*

- Applied **causal inference frameworks** to estimate LifeSync's impact, revealing a **3–5% uplift** in advisor bookings.
- Estimated **heterogeneous treatment effects** via **CATE models** to identify high-response customer subgroups.
- Formulated **policy optimization** strategies achieving up to **15% higher engagement** in top-tier segments.
- Demonstrated potential **2–3× efficiency gains** through targeted causal policy evaluation and ROI simulation.

Automated Microscopic Phenotyping for Arabidopsis Seeds **Sept 2022 – Apr 2023**
Guide: Dr. R Baskar, Developmental Genetics Laboratory — Biotechnology Department *IIT Madras, Chennai, India*

- Developed an **automated imaging pipeline** for phenotypic classification of **Arabidopsis** seed lines in **MeioSeed**.
- Implemented **OpenCV**-based segmentation and **watershed clustering**, achieving **94.5% accuracy** (AUC: 0.75).
- Engineered a scalable **Python automation tool** for high-throughput seed counting and phenotype labeling.

Scholastic Achievements

- **Secured a rank in the top 0.7%** among the 1 million students across India, in **JEE Advanced, 2021**
- **Cleared the** Pre-Regional Maths Olympiad **(PRMO)** and Regional Maths Olympiad **(RMO)**(2019)
- **Among the 28 students out of 250+ applicants** to get shortlisted for the **Young Research Fellowship**
- **Won bronze medals in Inter IIT Tech Meets 11.0 & 12.0 in Gen-AI and Multimodal-AI competitions.**
- **Finished 1st among 50+ teams** in a Hackathon involving potholes detection using Computer Vision
- Participated in an **Industrial AI** 24-hr hackathon conducted by **Temenos** and finished **4th** with over 200+ teams

Key Technical Projects

AMR-MoEGA: Hybrid Framework for Antimicrobial Resistance Evolution **Jan 2024 – May 2024**
Guide: Dr. Manikandan Narayanan, Bioinformatics and Integrative Data Science Lab *IIT Madras, Chennai, India*

- Developed a framework combining **Genetic Algorithms** with a **MoE classifier** for AMR genotype prediction.
- Processed genomic data via **BWA, SAMtools, VCFtools** to extract SNP features for ML-driven fitness evaluation.
- Simulated adaptive evolution with **HGT-based crossover**, achieving convergence toward high-resistance genotypes.
- Achieved robust genotype prediction with MoE classifier (**Accuracy: 93.4%, MCC: 0.87, AUC-ROC: 0.95**).

RL-Driven Dynamic Flux Balance Analysis for Microbial Metabolism **Jan 2024 – May 2024**
Guide: Dr. Karthik Raman, Computational Systems Biology Lab *IIT Madras, Chennai, India*

- Integrated **multi-agent reinforcement learning** with **dFBA** to optimize gene regulation and metabolic fluxes.
- Applied **genetic algorithms** for genotype space exploration, identifying high-impact regulatory strategies.
- Quantitatively evaluated metabolic adaptation, elucidating interplay between regulation and community dynamics.

Single-Cell RNA-seq Analysis for Therapeutic Target Discovery in OSF **Jan 2024 – May 2024**
Guide: Dr. Meiyappan Lakshmanan, Systems Biotechnology and Cellular Engineering Lab *IIT Madras, Chennai, India*

- Performed **scRNA-seq analysis with Seurat** to delineate diverse cell populations in OSF tissue samples.
- Applied **UMAP clustering and differential gene expression** to quantify cellular heterogeneity.
- Identified **key cell types and DEGs**, revealing potential biomarkers and therapeutic targets in OSF pathogenesis.

Enhancing Privacy-Utility Trade-offs in DP-Adam via Correlated Noise **Jul 2024 – Nov 2024**
Guide: Dr. Krishna Pillutla, Department of Data Science and Artificial Intelligence (DSAI) *IIT Madras, Chennai, India*

- Implemented **BLT-correlated noise** in DP-Adam, reducing gradient variance by **12%**, while maintaining privacy.
- Analyzed the effect of **direct vs. separate moment-wise noise injection** on the estimators for faster convergence.
- Demonstrated improved model utility under stringent budgets ($\epsilon \leq 2$) with up to **8% accuracy gain** over baseline.
- Extended **correlated noise mechanisms** from SGD to adaptive optimizers, establishing a foundational framework.

Fashionly.AI: Multimodal Multi-purpose AI-Powered Fashion Assistant **Jan 2025 – May 2025**

- Developed a fashion assistant leveraging **CLIP-ViT** for image classification and **Gemini LLM** for styling advice.
- Implemented a **digital closet** with tagging, seasonal filters, outfits, and usage analytics using **Flask endpoints**.
- Designed **Flask REST APIs** and secure session management for wardrobe, outfit creation, and AI chat interaction.
- Built real-time e-commerce scraping with **cosine similarity** over CLIP embeddings for product recommendations.
- Developed an interactive UI with outfit composer, uploader, and AI chat rendering, enabling multimodal workflows.

Relevant Coursework

- **Artificial Intelligence & Machine Learning** Foundation of Machine Learning, Introduction to Deep Learning, Modern Computer Vision, Advanced Topics in Artificial Intelligence, Machine Learning Operations Lab, Differential Privacy in AI, Recent Advancements in Generative AI
- **Computational & Systems Biology** Computational Systems Biology, Bioinformatics, Computational Biology Laboratory, Protein Interactions: Computational Techniques, Computer Simulations of Biomolecular Systems, Computational Neuroscience, Analysis and Interpretation of Biological Data, Biostatistics
- **Mathematics & Statistics**: Functions of Several Variables, Series and Matrices, Linear Algebra for Engineers, Probability, Statistics and Stochastic Processes, Mathematical Foundations of Data Science, Statistical Inference
- **Algorithms & Computing**: Algorithms in Computational Geometry, Algorithmic Approaches to Computational Biology, Signals and Systems, Data Analytics Laboratory

Technical Skills

- **Languages**: Python, Matlab, Java, C, HTML, CSS, JavaScript, LaTeX, R
- **Libraries**: PyTorch, Tensorflow, Scikit-learn, OpenCV, Numpy, Pandas, Nltk, Optuna, HuggingFace, Streamlit
- **Tools**: Grafana, Prometheus, JupyterLabs, Docker
- **Proficiency**: Git Version Control, Data Analysis, Feature Engineering, Image Processing, Machine Learning, Deep Learning, Explainable AI, Generative AI, Prompt Engineering, Large Language Models, ML Deployment

Leadership Experience

Strategist, Artificial Intelligence Club — Center for Innovation, IIT Madras **Apr 2023 – Mar 2024**

- Led a cross-functional team of **50+ students** to design AI-driven solutions addressing real-world challenges.
- Collaborated with **industry, startups, NGOs, and faculty** to translate research concepts into applied AI projects.
- Organized and delivered **technical workshops** on core areas of Computer Vision, Deep Learning, and Applied AI.

Event Lead, Generative AI Workshop — Shaastra, Technical fest, IIT Madras **Feb 2023 – Mar 2023**

- Led and mentored a team conducting a hands-on workshop on **Generative Modeling** with over **100+** participants.
- Designed technical content covering **Autoencoders, GANs, CycleGANs, DCGANs**, and **Game Theory**.
- Facilitated conceptual coding sessions bridging theoretical intuition with practical implementation in generative AI.

Teaching Assistant, Programming and Data Structures, DSAI, IIT Madras **Jul 2025 – Nov 2025**

- Managed **Gradescope**: configured autograders, designed assignments, and oversaw automated evaluation pipelines.
- Supported grading, paper evaluation, and student mentorship, ensuring consistent feedback and course operations.
- Created interactive **Slido quizzes** and **hands-on** coding sessions to teach DSA in a more engaging and intuitive way.

Extracurricular

- **Won a Silver medal** in the Science and IT Quiz conducted by **IISc**
- Participated in Technological Quizzes conducted by **TCS, Times NIE, and KISA**
- One among the **15 students**, out of 200+ students to get selected for **National Sports Organization, Badminton**
- Secured the **Gold Medal** in the **inter-department badminton league**, representing my department.

Note: Currently in ninth semester (November 2025)