

Diya Goswami

+91-6909371878 | diya13goswami@gmail.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

Vellore Institute of Technology, Bhopal

B. Tech Computer Science & Engineering with specialization in Health Informatics
8.98 CGPA

Bhopal, India

Sep 2022 – July 2026

Auxilium Girls' School

Higher Secondary
95.2 Percentage

Tripura, India

March 2021 – April 2022

Auxilium Girls' School

Higher School
96.2 Percentage

Tripura, India

March 2019 – April 2020

PROJECTS

EchoRetail: Retail Feedback Captured and Analyzed by AI

Aug 2025-Ongoing

- An AI-driven retail analytics system that generates synthetic transaction datasets using GANs and simulates realistic customer behavior.
- Implemented LLM-powered feedback analysis with ChromaDB and Gemini embeddings, enabling natural language queries on customer sentiment and trends.
- Applied topic modeling with BERTopic to cluster feedback into themes and visualized evolving trends in customer preferences and complaints.
- Conducted exploratory data analysis (EDA) on synthetic datasets, uncovering revenue patterns, category performance, and customer segmentation.

AI-Augmented Cardiac Risk Prediction using Synthetic Data Generation Techniques

Oct 2024-May 2025

- Built a synthetic data pipeline using CTGAN, VAE, and Table Diffusion to address dataset imbalance in heart disease prediction.
- Trained multiple models (SVM, XGBoost, CNN, Ensemble) achieving 85.85% accuracy with KNN on VAE data, 83.7% with Ridge Regression on Diffusion data, and consistent improvements across precision/recall (avg. 0.86 F1-score).
- Validated results with confusion matrices, showing reduced false negatives and improved diagnostic reliability using synthetic datasets.

SkinSight: Intelligent Skin Type Detection System

June 2023- April 2024

- Built a real-time CNN, ResNet-50 and Haar Cascade-based skin type detector on Raspberry Pi 5 + Logitech C920, achieving 80.38% accuracy (Dry: 56/73, Normal: 98/120, Oily: 100/123).
- Optimized with TensorFlow, Keras, OpenCV, and quantized models for low-latency edge deployment; integrated LED ring light and one-way mirror UI for clinical-grade usability.
- Applications in personalized dermatology, cosmetic recommendations, and telemedicine; demonstrated at ETESM-2025.

TECHNICAL SKILLS

Languages: Java, Python, C++, SQL

Frameworks and Libraries: TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, Keras, Matplotlib

Generative AI & LLMs: LLMs (Gemini, GPT), RAG, LangChain, LangGraph, ChromaDB, PineconeDB

Developer Tools: Git, GitHub, Jupyter, Google Colab, Tableau,

Other Skills: Data Structures, Problem Solving, Machine Learning, Deep Learning, Generative AI

CO-CURRICULAR

- Smart India Hackathon 2024 Finalist
- Health Hackathon JHU & VITB Finalist
- Presented and Published research work on ML in cardiac disease prediction in ICDCC 2024
- Core Member, Eureka Club, VIT Bhopal