Diya Goswami

CONTACT

diya13goswami@gmail.com
+91- 6909371878
Agartala, India
LinkedIn
Github

EDUCATION

B. Tech
Computer Science and Engineering with specialization in Health Informatics
Vellore Institute of Technology
Bhopal
September 2022 - July 2026
8.98 CGPA

Higher Secondary
Auxilium Girls' School
March 2021- April 2022
95.2 Percentage

High School
Auxilium Girls' School
March 2019- April 2020
96.2 Percentage

CO-CURRICULARS

- Smart India Hackathon 2024 Finalist
- Health Hackathon JHU & VITB Finalist
- Presented and Published research work on ML in cardiac disease prediction in ICDCC 2024
- Presentation and Publication of SkinSight in ETESM 2025
- Core Member, Eureka Club, VIT Bhopal
- Open-Source Contributor, GitHub

PROJECTS

EchoRetail: Retail Feedback Captured and Analyzed by AI
August 2025- October 2025

- Designed and trained a GAN-based data synthesis engine (PyTorch, 500 epochs) that generated 10,000 privacy-safe retail transactions, enabling scalable analytics and reducing manual data collection costs.
- Implemented a Retrieval-Augmented Generation (RAG) system using Google Gemini, ChromaDB, and LangChain, enabling natural language insights from more than 10 thousand customer reviews within seconds.
- Built a comprehensive NLP analytics suite featuring aspect-based sentiment analysis, BERTopic-driven theme discovery, and temporal trend visualization, providing actionable customer intelligence.

CardiaSynth: Multi-Model Synthetic Data Generation for Improved Diagnostics
November 2024-May 2025

- Engineered a multi-model synthetic data pipeline leveraging CTGAN, VAE, and Table Diffusion, generating balanced cardiac datasets that boosted ML accuracy by 15–20% across five algorithms.
- Developed an ensemble learning framework (SVM, XGBoost, CNN, Ridge, KNN) achieving 85.85% accuracy and 0.86 F1-score, validated across three synthetic datasets.
- Optimized confusion matrices to reduce false negatives by 25%, enhancing clinical reliability in cardiac risk prediction.

SkinSight: Intelligent Skin Type Detection System
June 2023- April 2024

- Developed production-ready skin type detection system combining CNN, ResNet-50, and Haar Cascade models on Raspberry Pi 5, achieving 80.38% accuracy with consistent 76-81% per skin type performance in real-time edge inference
- Designed a clinically adaptive interface integrating LED illumination, one-way mirror display, and TensorFlow Lite quantization, ensuring sub-second predictions on constrained hardware.
- Architected clinical-grade user interface with hardware-software integration including LED ring light for controlled illumination, one-way mirror display, and TensorFlow Lite model quantization enabling sub-second inference latency on resource-constrained edge devices

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

CERTIFICATIONS

- Google Data Analytics
- FutureSkills Generative AI Fluency
- Oracle Java Foundations
- Career Essentials in Data Analysis by Microsoft and LinkedIn