ANURAG INDORA

anuragindora474@gmail.com | +91-7056717349 | linkedin.com/in/anurag-indora |
github.com/AnuragIndora | dev-ai-09.vercel.app

Education

Indian Institute of Technology (ISM), Dhanbad *May 2025*
Integrated M.Tech in Mathematics & Computing

- Coursework: Machine Learning, Deep Learning, Optimization, Statistics, Linear Algebra, Data Structures.
- Secured top academic performance in Mathematics & Computing cohort.

Experience

Machine Learning Intern — CodSoft *May 2024 – June 2024*

- Designed and deployed a **fraud detection system** using SMOTE, ADASYN, and XGBoost, achieving **97.2% accuracy** on **284K+ transactions**.
- Reduced false positives by **23%**, improving precision and business decision quality.
- Built modular ML pipelines (loader, transformer, trainer, evaluator) enabling scalable and reusable workflows.
- Integrated **CI/CD automation (GitHub Actions)** for continuous training & deployment, reducing downtime by 35%.

Projects

Transformer-Based Text Summarizer — *Python, Hugging Face, BART, FastAPI, Streamlit*
GitHub Repo

- Developed an **abstractive text summarization pipeline** fine-tuned on **50K+ CNN/DailyMail samples** using `facebook/bart-base`.
- Implemented a 6-stage ML pipeline (ingestion→validation→training→evaluation→prediction) improving iteration speed by 40%.
- Achieved **ROUGE-L 20.67** and **3.6 eval/s** using FP16 mixed precision on a 4GB GPU.
- Deployed **FastAPI inference service** with Streamlit UI for real-time summarization and LLM context compression.

Movie Recommendation System — *Python, TF-IDF, Cosine Similarity, TMDb API, Streamlit*
GitHub Repo

- Built a content-based recommender for **5K+ movie titles** leveraging TF-IDF and cosine similarity.
- Reduced model footprint by **80%** using sparse matrix storage.
- Integrated TMDb API to automatically update metadata, ensuring real-time relevance for 5K+ movie titles, and reduced manual data entry errors by 8 %, improving the quality of recommendations.

Research & Publications

Master's Thesis (2025) — *Facial Expression Recognition in Children using Graph Neural Networks*

- Designed a GNN-based emotion recognition pipeline with attention-based feature aggregation and graph convolution.
- Explored representation learning for low-data affective computing scenarios.

Achievements

- **Silver Medal (Top 50 Rank)** — NeurIPS Open Polymer Challenge 2025 (Kaggle).
- Published research in computer vision and graph learning applications.
- Active contributor to Kaggle and open-source ML repositories.

Technical Skills

Languages:	Python, C++, Java, SQL, Bash
Frameworks:	PyTorch, TensorFlow, scikit-learn, Hugging Face, FastAPI, Streamlit, XGBoost
Tools:	Docker, Git, GitHub Actions, Makefile, Azure, CI/CD Pipelines, API Deployment
Data & Modeling:	Feature Engineering, GPU Acceleration, Scalable Algorithms, SQL Databases (MySQL)

Additional Strengths

- Collaborative experience integrating research prototypes into production ML systems.
- Strong communication for technical documentation and presentations.