# Ayush Ranjan

Ranchi, Jharkhand, India | +91 9142115113 | ayushranjan1271@gmail.com
linkedin.com/in/ayush-ranjan-1b0a47300 | github.com/Ayush-1271 | ayush-1271.github.io

## Summary

Machine Learning Engineer and NLP Researcher pursuing an Integrated M.Tech in Computer Science (Computational and Data Science) at VIT Bhopal (2023–2028). Published researcher in a Q1 Scopus journal (CiteScore 8.3) on cryptographic image security. Experienced in LLM post-training (RLHF, HHH alignment), NLP pipelines, computer vision, and full-stack ML deployment. Solved 300+ DSA problems on LeetCode and am an active open-source contributor. Seeking roles in machine learning engineering, NLP, or AI research.

## Technical Skills

**Languages:** Python, Java, C++, TypeScript, JavaScript, SQL, HTML/CSS

**Machine Learning / AI:** PyTorch, TensorFlow, Scikit-learn, XGBoost, Random Forest, LLM Fine-tuning, RLHF, Post-training, Prompt Engineering, HHH Alignment

**NLP:** Transformers, BERT, LSTM, TF-IDF, spaCy, NLTK, sentence-transformers, Hugging Face

**Data Science:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Feature Engineering, Model Evaluation

**Cryptography & Security:** LFSR, GF(3) Fields, Image Encryption, Chaos Theory, NPCR/UACI Analysis

**Frameworks & Tools:** FastAPI, Streamlit, Next.js, Docker, Git, AWS, Selenium, BeautifulSoup

**Databases:** MongoDB, ChromaDB, SQL

**Core CS:** Data Structures & Algorithms, Object-Oriented Programming, DBMS, Operating Systems, Networking

## Experience

**Ethara.ai** **Feb 2026 – Apr 2026**
*LLM Post Trainer Intern* *Remote*

- Scored and ranked 500+ LLM-generated responses across helpfulness, honesty, and harmlessness (HHH) dimensions using RLHF, directly improving model alignment quality.
- Curated 150+ adversarial prompts and fine-tuning examples to expose failure modes and reduce harmful outputs in post-training pipelines.
- Contributed preference data that fed iterative fine-tuning cycles, accelerating model improvement along 3 evaluation dimensions.

**Freelancing Club, VIT Bhopal** **Jan 2025 – Present**
*Core Member, Machine Learning* *Bhopal, India*

- Delivered 5+ client ML projects end-to-end — from data collection through model deployment — covering classification, regression, and automation use cases.
- Engineered 3 automated data analysis pipelines that slashed manual processing time by 60%, freeing clients to focus on decisions rather than data wrangling.

**Codetech** **Nov 2024 – Jan 2025**
*Data Science Intern* *Remote*

- Boosted predictive model accuracy by 12% through systematic feature engineering, hyperparameter tuning, and cross-validation across 4 datasets.
- Cut model training time by 35% via pipeline bottleneck refactoring and early-stopping strategies, accelerating the overall iteration cycle.

## Projects

### NoteMind — Multi-Agent AI Memory System

*Tech: Next.js, FastAPI, ChromaDB, MongoDB, Gemini 2.5*

- Architected a persistent semantic knowledge graph for AI agent swarms; each agent writes atomic notes into ChromaDB, instantly retrievable by all agents via semantic similarity.
- Eliminated O(n2) token scaling — a 5-agent swarm maintains O(k) cost regardless of session length, removing the context-limit ceiling entirely.

### MangaLens — Chrome Extension, v1.0.8 (Released)

*Tech: JavaScript, Chrome MV3, Gemini Vision, Canvas API*

- Shipped a production Chrome extension that uses Gemini Vision to detect speech-bubble bounding boxes in manga images, fills bubble backgrounds, and overlays translated text in 18 languages.
- Extended compatibility to lazy-loaded and infinite-scroll pages (MangaDex, Webtoon, MangaPlus); translation memory preserves character name consistency over 100+ page sessions.

### MarketLens — Probabilistic Stock Forecaster

*Tech: PyTorch, CNN, Transformers, RAPID Framework*

- Constructed a CNN + Transformer ensemble that outputs P10/P50/P90 confidence bands rather than a single point prediction, rendering uncertainty quantification explicit.
- Designed the RAPID trust-scoring framework (5 dimensions) that objectively flags when the model's forecast should not be acted on — built-in, not bolted on.

### Legal NLP Classifier

*Tech: Python, TF-IDF, SVM, Scikit-learn*

- Achieved 93% classification accuracy on 5,000+ Supreme Court of India judgments spanning 5 legal categories (constitutional, criminal, civil, tax, family law) using TF-IDF and SVM.
- Surfaced keyword evidence behind each classification decision, ensuring model reasoning is transparent and auditable for non-technical legal staff.

## Research & Publications

### A New Approach for Image Security Enhancement Using Ternary Logic LFSR for Cryptographic Applications

*Advances in Decision Sciences*, Vol. 30(3), 2026, pp. 184–214. Q1 Scopus (CiteScore 8.3). DOI: 10.47654/v30y2026i3p184-214

Proposed a novel GF(3) LFSR-based stream cipher for medical image encryption. Achieved 98.04% NPCR and 6.80-bit Shannon entropy. Produces an exponentially larger keyspace than binary LFSR baselines.

## Education

**Vellore Institute of Technology, Bhopal** — **2023 – 2028**

*Integrated M.Tech in Computer Science — Computational and Data Science* — *Bhopal, India*

## Certifications

**Google IT Support Professional Certificate** — Coursera / Google Career Certificates, Feb 2026 (Verified, Credly)

**AWS Cloud Practitioner Essentials** — Amazon Web Services

**Deloitte Data Analytics Virtual Simulation** — Forage

**Python Programming Certificate**; **Java Programming Certificate**

## Achievements & Open Source

**Competitive Programming:** 300+ DSA problems solved on LeetCode; Codeforces Pupil rank (handle: Raj_ay). Core topics: Dynamic Programming, Graphs, Trees.

**Open Source:** Contributor to SunPy (solar physics Python library); participant in Social Summer of Code (SSOC 2025).

**Community:** NSS Volunteer — co-organized a blood donation drive (400+ units collected) and a plantation drive (10,000+ trees).