

Ankan Ghosh

LinkedIn: [linkedin.com/in/ghoshankan](https://www.linkedin.com/in/ghoshankan)

GitHub: github.com/ankan00v

Portfolio: ankanghosh-portfolio.app

Email: ghoshankan005@gmail.com

Mobile: +91-9046359825

SKILLS

Languages: Python, SQL, C++

Data Analysis & Visualization: Pandas, NumPy, SciPy, EDA, Power BI, Excel

AI/ML: Scikit-learn, NLP, Recommendation Systems, A/B Testing, Model Evaluation

Backend: FastAPI, REST APIs, Data Pipelines

Databases: PostgreSQL, MySQL, MongoDB

Tools & Platforms: Git, Docker, Jupyter, Google Colab

Soft Skills: Problem Solving, Analytical Thinking, Collaboration, Adaptability, Communication

INTERNSHIP

SDE Intern | NIIT Limited

June'26 – Aug'26

- Developing LLM-powered backend features and AI/ML pipelines using FastAPI; collaborating across engineering and data science to deliver scalable AI-powered services per sprint.
- Participating in code reviews, sprint planning, and architecture discussions; documenting technical specifications and implementation decisions.

IT Analyst Intern | Eastern Coalfields Limited (Govt. of India) | [Link](#)

Jul'25 – Aug'25

- Engineered Python ETL pipeline to automate SAP data extraction and report generation across 2 departments, eliminating 3 hours/day of manual processing and reducing data errors from 12% to under 2%.
- Streamlined Power BI reporting by automating pipeline refresh via REST API and batch scripts, reducing delivery time from 24 hours to 1 hour and removing 4 daily manual interventions.

PROJECTS

VidyaVerse | Python, SQL, NLP, Vector Search, Recommendation Systems, A/B Testing | [GitHub](#)

Present

- Engineered an opportunity discovery platform processing 1700+ postings from 200+ sources with 384D vector search, Sentence Transformers, and multi-mode ranking; fine-tuned JobBERT, improving skill-extraction F1 from 0.180 to 0.601.
- Strengthened ML reliability across Python and PostgreSQL by relabelling 8,374 records and implementing automated validation across 502 tests and 17 CI/CD gates, preventing unvalidated model promotion.

LPU Smart Campus | Python, FastAPI, PostgreSQL, MongoDB, Data Pipelines | [GitHub](#)

Jan'26 – Apr'26

- Architected a real-time data platform for campus analytics and attendance insights, enabling automated attendance tracking and decision workflows across 14+ modules.
- Developed scalable data pipelines integrating 200+ APIs and SQL-based data processing (PostgreSQL), supporting high-frequency data ingestion and analytics.

Public Policy Analytics | Python, Pandas, NumPy, Excel, Data Visualization | [GitHub](#)

Apr'25 – May'25

- Conducted large-scale analysis on 14K+ government welfare records (IGNDPS) to identify demographic trends, regional gaps, and digital inclusion drivers.
- Designed interactive Excel and Power BI dashboards with KPI tracking, correlation analysis ($r \approx 0.83$), and time-series insights to support data-driven policy evaluation.

CERTIFICATIONS

- Oracle Cloud Infrastructure 2025 Certified Data Science Professional | [Oracle](#) Jul'25
- Oracle Cloud Infrastructure 2025 Certified Generative AI Professional | [Oracle](#) Jul'25
- Oracle Autonomous Database Cloud Certified Professional | [Oracle](#) Sep'25

ACHIEVEMENTS

- AIR 122 — Naukri Campus Young Turks 2025. Top 500 / 500,000+ participants nationally (Coding, DS & AI track).
- AlgoUniversity Accelerator Graph Camp - Top 50 nationally among 40,000+; ranked 2nd at LPU.
- Dean's List recipient, Lovely Professional University. Top 10% of batch by academic performance.
- GSSoC 2026 Contributor — merged 4 PRs to OpenCode, implementing Discord OAuth 2.0, session management and Redis caching, reducing latency by ~200ms via; 85%+ test coverage maintained.

EDUCATION

Lovely Professional University

Bachelor of Technology

Computer Science and Engineering; CGPA: 8.16

Phagwara, Punjab

Aug' 23 – Present