CHIRAG KALUCHA

+918699714767 • chiragkalucha15@gmail.com • LinkedIn • GitHub

EDUCATION

Chandigarh University, Mohali, Punjab *2022-2026*

Bachelor of Engineering in Computer Science (specialization in AI and Machine Learning) *CGPA:8.7/10*

Relevant Coursework: Big Data Analytics, Data Mining, Predictive Analytics, Machine Learning, AI, DBMS, Soft computing.

TECHNICAL SKILLS

Programming & Software Engineering: Python (Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Faker, Streamlit), SQL (PostgreSQL, MySQL, Window Functions, CTEs, Query Optimization, Joins), REST APIs, FastAPI, JSON, API Testing & Integration.
Data Engineering & Infrastructure: Apache Kafka, Snowflake, AWS Redshift, Docker, ETL Pipelines, Data Preprocessing, Data Analytics.
Machine Learning & AI: Scikit-Learn (Random Forest, XGBoost, Logistic Regression, Ensemble Methods), Generative AI (RAG Architectures, Vector Databases, Prompt Engineering), Model Evaluation (Precision, Recall, F1), Supervised Learning, Churn Prediction.
Analytics & Visualization: Funnel Analysis, Cohort Analysis, A/B Testing, Hypothesis Testing, Statistical Inference, KPI Design, Power BI, Excel (Power Query, Pivot Tables, Advanced Formulas), Heatmap Visualization, Sales Forecasting.
MLOps & Deployment: Docker, CI/CD Pipelines, GitHub Actions, Model Deployment.

PROFESSIONAL EXPERIENCE

Ministry of Home Affairs, Delhi *June 2025 – December 2025*

AI and Data Analyst Intern

- Reduced onboarding time for oceanographic analysts from 3-4 weeks to <8 hours by developing an interactive acoustic simulator deployed on the department's intranet. Enabled hands-on learning of domain concepts and sensor technologies across 5 departmental workstations.
- Integrated Hugging Face transformer models into the simulator to enable real-time spectrogram interpretation, providing automated acoustic pattern recognition that eliminated manual analysis bottlenecks for junior researchers.

Internsellite, Sohna, India *Jun 2024 - Aug 2024*

Data Science Intern

- Identified $15000+ in potential monthly revenue optimization opportunities by analysing 20,000+ Amazon sales transactions using SQL and Python, discovering geographic clusters with 30%+ higher customer satisfaction rates for targeted product placement strategies.
- Built automated data preprocessing pipelines in Python (Pandas, NumPy) that reduced manual data cleaning effort by approximately 60%, enabling faster iteration on analytical insights for business stakeholders.

PROJECTS

E-Commerce Funnel Analysis & Churn Prediction System with Automated Engagement Optimization

(Python, PostgreSQL, Faker, Scikit-Learn, XGBoost, Docker, Streamlit, Cohort Analysis, A/B Testing)

- Architected an end-to-end analytics pipeline to reduce customer drop-off rates by 22% across the conversion funnel by engineering a synthetic e-commerce dataset (500K+ user sessions) with Python Faker, modelling realistic user journeys through registration, browsing, cart addition, and checkout stages.
- Developed a multi-model churn prediction system using Random Forest and XGBoost achieving 89% recall on high-risk user identification, enabling proactive intervention campaigns targeting users with >70% churn probability before session abandonment.
- Dockerized the entire stack (PostgreSQL database, ML models, Streamlit dashboard) and deployed on GitHub with CI/CD pipelines for reproducible deployment, reducing setup time from days to <15 minutes for new analysts joining the project.

Oceanographic Geospatial Analytics & Sensor Deployment Optimization System

(Python, GeoPandas, Shapely, Folium, Streamlit, K-Nearest Neighbor)

- Processed and analysed 5GB+ of geospatial CTD (Conductivity-Temperature-Depth) sensor data covering 4 million+ coordinate points across the Indian Ocean to identify critical data gaps and prioritize sensor deployment zones for oceanographic research authorities.
- Identified 6 distinct deviation categories between ground-truth bathymetry and open-source aggregated datasets, revealing ±15-30% variance in the western Indian Ocean (East Africa corridor) and Indonesian archipelago belt.
- Built an interactive Streamlit dashboard with real-time heatmap visualizations using Folium, enabling stakeholders to explore data density patterns across regional grids and make data-driven decisions on exploration priorities; deployed for daily operational use by government research teams.

ACHIEVEMENTS AND RECOGNITION

- **Smart India Hackathon 2024 National Finalist** – Selected among top 0.5% of 15,000+ participating teams nationwide for government-sponsored innovation challenge addressing real-world societal problems.
- Published *3 peer-reviewed* research papers: "**Multi-Agent AI Framework for Developer Assistance**" (*ICDAM Conference*), "**Comparative Evaluation of ML Algorithms for Bank Fraud Detection**" (*Wiley Library*), and filed patent for "Virtually Aided Cashing Machine" (*Patent Pending*).
- **Department Representative** – Career Development & Training, Chandigarh University
- **Chief Executive** – Events & Programs, CAC Club
- **Coursera Campus Ambassador** coordinating 500+ students for skill development initiatives.
- Certifications: Python for Data Science (*IBM*), Applied Machine Learning in Python (*University of Michigan*), Generative AI for Data Analytics (*Microsoft)* via Coursera.