Ravi Kumar

B.Sc.(Hons.) Data Science & Artificial Intelligence, IIT Guwahati

Roll No: [redacted] | CPI: 7.3 / 10

[redacted]

[redacted]

GitHub · LinkedIn · Portfolio · Kaggle · Codolio

Professional Summary

Data Science & AI undergraduate at **IIT Guwahati**, currently driving impact as **AI Product Engineer Intern at PhysicsWallah (PW)**. Specializing in AI/ML, LLM integration, Generative AI, workflow automation, and business intelligence dashboards. Experienced in building end-to-end ML & GenAI pipelines using Python, FastAPI, Hugging Face, and Looker Studio. Hackathon winner, Kaggle-certified SQL practitioner, and competitive programmer with **600+ DSA problems solved** and global rank under 1500 on Codolio.

Education

Degree	Institute	CPI / %	Year
B.Sc.(Hons.) of Data Science & AI	Indian Institute of Technology (IIT) Guwahati	7.3 / 10	2023 – 2027
Senior Secondary (Class XII)	Bihar School Examination Board (BSEB)	85%	2022

Experience

PW (PhysicsWallah)April – June 2026, AI Product Engineer Intern, PWIOI Team (Noida, Uttar Pradesh, India). Built **15+ Looker Studio dashboards**(live link) for admissions, sales & funnel tracking, used weekly by leadership – PWIOI Sales Funnel Dashboard. Developed a **Streamlit automation tool** (live link) to extract 200+ candidate videos from Mettle PDFs, boosting hiring efficiency. Created **7+ Python/Apps Script automations** (score mapping, ERP sync, scholarship tracking) eliminating hours of manual work. Built a **Telegram bot** for marketing automation and authored **3 BRDs** (Loan, Bonafide, Refund modules) for product engineering teams.

Technical Skills

Languages: Python, SQL, C++, JavaScript **DSA:** 600+ problems solved · Global Rank <1500 on Codolio

ML / DL: XGBoost, LSTM, Neural Networks, Random Forest, Gradient Boosting, Regression, Classification, Feature Engineering

NLP / GenAI: RAG, TF-IDF, Vector Embeddings, Transformers, Prompt Engineering, LLM Integration, Cosine Similarity

Data & BI: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, EDA, Looker Studio, Power BI, Funnel Analytics

Databases: MySQL, PostgreSQL, CTEs, Window Functions, Joins, Query Optimization **Automation:** Python Scripting, Telegram Bots, Data Mapping

Tools: Git, GitHub, Streamlit, FastAPI, Docker, Alembic, Jupyter, Google Colab, VS Code

Projects

Admission Samadhan & Sales Funnel Dashboard | *Looker Studio, Google Sheets, Python, SQL* **[Dashboard]** Built a high-impact BI dashboard for the PWIOI team to track the 2026 registration funnel and admission samadhan, with Python-based data sync from Google Sheets. Automated daily reporting with real-time visibility into sales metrics and student queries for leadership; integrated complex city-wise and funnel-stage filters with a clean, user-centric UI.

SpaceIQ — AI-Powered Space Booking Platform | *Python, Streamlit, LLM Integration, FastAPI* **[GitHub]** Developed a production-style office space booking platform featuring AI-driven recommendations, smart search and filtering, and a demand-based dynamic pricing signal engine to optimize space utilization. Integrated a chat-assisted booking interface powered by an LLM, real-time booking analytics dashboard, and a FastAPI backend for session management — backed by a self-contained mock dataset for demo-ready deployment.

Candidate Evaluation System for Hiring Pipeline | *Python, FastAPI, PostgreSQL, Alembic, LLM, GitHub API, Docker* **[GitHub]** Built a production-grade, multi-stage hiring pipeline with FastAPI backend and PostgreSQL (via Alembic migrations): Stage 1 performs rule-based completeness checks; Stage 2 uses an LLM to score written responses on a 0–5 rubric and flags AI-generated content; Stage 3 fetches GitHub activity signals and runs LLM-powered code quality review. Composite scoring engine automatically classifies candidates as *reject / borderline / invite / priority* with human-reviewer override. Entire system containerized using Docker Compose for one-command deployment.

AegisMarket: Autonomous Financial Agent Stress-Testing Platform | *Python, Flask, AI Agents, Pandas, NumPy, SQL* **[GitHub]** Designed a financial market simulator that models agent behavior and price dynamics on a synthetic market, enabling micro-economic stress-scenario testing (inflation shocks, liquidity crunches). Built an end-to-end stress-testing pipeline using Python, Flask, and Pandas/NumPy for data simulation, analysis, and visualization of agent-driven market outcomes.

Walmart Sales Data Analysis | *Advanced SQL, MySQL, CTEs, Window Functions, Power BI* **[GitHub]** Analyzed a large-scale Walmart retail dataset using advanced MySQL — modular CTEs for complex query logic, window functions for running totals and branch-level rankings, multi-table joins for cross-branch performance comparisons. Uncovered top-performing product categories, seasonal sales spikes, and underperforming branches; translated findings into interactive Power BI dashboards tailored for non-technical business stakeholders.

Achievements & Certifications

Innovation Hackathon — 3rd Place, PhysicsWallah (2024): Secured 3rd place among 500+ participants across engineering colleges by building an AI-driven solution.

Competitive Programmer — Codolio: Solved **600+ DSA problems**; Global Rank under 1500. **Top 20%** — Kaggle ML Competition for applied ML & feature engineering.

Advanced SQL Certification — Kaggle (2025): CTEs, window functions, joins, analytical SQL workflows. **Certificate of Participation** — Codemet & Trilytics'24 Analytics Case Competition.