

Aryan Rajpurkar

 av.rajpurkar@gmail.com  +917718950601  Mumbai, India  <https://aryanrajpurkar.tech>  <https://github.com/aryan4codes>
 linkedin.com/in/aryan-rajpurkar

Education

B.Tech in Computer Science and Engineering (Data Science) - 9.2 CGPA, D.J. Sanghvi College of Engineering

2022 – 2026

Professional Experience

Data Engineering Intern, VisaFriendly ↗

06/2025 – Present

India

- Developed **real-time ETL pipelines** processing **100K+ daily job entries**, improving **accuracy by 28%** and reducing **processing time by 40%**
- Built **AI agents for resume and cover-letter generation** with **Crewai and n8n**, automating ~8K documents monthly and improving **interview response rates by 22%**
- **Technical Skills:** AWS (ECS, ECR, CloudWatch,RDS), Python, SQL, Crewai, n8n, ETL Pipelines, Data Modeling, LLMs

SDE (ML) Intern, Mantrika.ai ↗

01/2025 – 08/2025

India

- Led development of AI-powered HR recruitment platform as founding engineer, **implementing Agentic AI integration** and database schemas with **40% improved latency ,reducing cost per candidate selection by 20%**.
- Established **end-to-end agentic system architecture** and led core platform development.
- **Technical Skills:** TypeScript, Next.js, Langchain, LangSmith MySQL, Supabase, AWS, CrewAI , Agentic AI

NLP (ML) Engineer Intern, AI4Education Labs ↗

06/2024 – 07/2024

Germany

- Development of an **AI-driven assessment platform** using **item response theory and knowledge graphs**, impacting **5,000+ students across 10+ schools** in India and Germany.
- Engineered end-to-end NLG pipeline with **95% accuracy in personalized content recommendations**, leveraging AWS infrastructure
- **Technical Skills:** Knowledge Representation, Item Response Theory Research , AWS (S3, EC2)

Software Engineer Intern, Drilldown ↗

03/2024 – 08/2024

India

- Designed and deployed **multimodal AI agents** on Telegram, WhatsApp, and Instagram, reducing audit turnaround time by **35% for FMCG clients**
- Developed an **AI-powered Influencer Marketing platform** using **Instagram Graph API and LangChain** to match brands with creators based on **engagement analytics**
- **Technical Skills:** Generative AI, LangChain, LLMs, Prompt Engineering, GCP (BigQuery, Cloud Run), SQL, AI Agents

Projects

Edudata, Intelligent Document Search and Management Platform ↗

- Manages and analyzes documents with **Precedent Matching**, machine-readable checks, and **automated approval workflows**.
- Features admin-user workflows, **real-time metadata generation**, and detailed **RAG intelligent search via embeddings (hybrid retrieval BM25 + vector, reranking)**.
- Implemented under the **Government of India's e-Office** Mode Project.
- **Technical Skills:** NextJS, Python (Flask), MongoDB (GridFS), Clerk API, BM25, OCR, Elasticsearch, Apache Kafka.

Equitas, Real Time AI Safety and Monitoring Platform ↗

- **Enterprise AI safety layer for LLMs** with multi-layer detection (toxicity, bias, hallucination, jailbreak) **using custom transformer models** (Detoxify, SHAP, LIME); no vendor lock-in. [Custom Python Package Link](#) ↗
- Features **real-time explainability with SHAP/LIME token-level analysis**, multi-tenant architecture, credit-based billing, and a drop-in LLM APIs replacement with safety metadata.
- **Technical Skills:** Python (FastAPI), React/TypeScript, MongoDB, Clerk API, PyTorch, Transformers, SHAP, LIME, Redis, Docker.

Leadership

Chairperson - DJS-S4DS, National Data Science Committee of India ↗

07/2024 – Present

- Designed and executed technical workshops and bootcamps, **mentored 200+ students** in machine learning, AI and data analytics

Awards

Winner at Smart India Hackathon (SIH) 2024, National Hackathon hosted by Government of India ↗

Selected Scholar, Amazon ML Summer School 2024, Selected out of 61,000 participants

Best Student Chair 2025, Society of Data Science - India ↗

Winner of 5 National Hackathons | Awards Worth ₹2+ Lakhs ↗

Skills

Languages: Python, Java, JavaScript, Typescript, SQL

Web Development Frameworks: Flask, React, NextJS , SpringBoot , RESTful API development

Cloud & Data Engineering: AWS (EC2, S3, Athena, Glue), GCP (Cloud Run) , Apache (Kafka, Spark, Airflow)

DevOps and Containerization: Git, GitHub, Docker, Kubernetes, CI/CD , Microservices

Data Science and Machine Learning: PowerBi, NLP, TensorFlow, Computer Vision, Generative AI, Federated Learning , Data Analytics, Graph DBs (neo4j) , FAISS , Pinecone , AgenticAI , VectorDB (Astra , Pinecone) , RAG , MCP