

Dev Pandya

8200553579 | b22ai016@iitj.ac.in | [linkedin.com/in/dev-pandya-6769a5252/](https://www.linkedin.com/in/dev-pandya-6769a5252/) | github.com/devpandya5804

EDUCATION

Indian Institute Of Technology Jodhpur

Bachelor of Technology in Artificial Intelligence and Data Science

Oct 2022 - May 2026

Puna International School

Class 12

Jun 2020 - Jun 2022

Zyodus School For Excellence

Class 10

Apr 2015 - Apr 2020

EXPERIENCE

Software Development Intern

United Techlab

May 2025 - Jul 2025

- Fine-tuned **Qwen 3 8B** LLM using **QLoRA** and **Unsloth**, reducing model size by **60%** via **GGUF** export for **Ollama**, enabling low-latency inference on local devices with limited compute
- Built and deployed a **RAG pipeline** on **AWS RDS** with **pgvector**, improving domain-specific semantic search accuracy by **40%** using keyword-weighted retrieval
- Developed scalable backend with **FastAPI**, **AWS EC2**, **AWS Lambda**, and **Redis** for cron-based GPU scheduling; implemented **CI/CD pipelines** using **GitHub Actions**
- Secured production deployment with **Nginx** reverse proxy, **HTTPS** via SSL certificates, and implemented **unit/integration testing** with detailed technical **documentation**
- Collaborated in an **Agile Scrum** environment, resolving **Git** merge conflicts, ensuring professional communication, and delivering features within sprint deadlines

Quantitative Research Consultant

WorldQuant

Oct 2023 - Jan 2025

- Engineered and evaluated **alpha factors** for US and China equity markets using 10+ years of historical price-volume data
- Increased Sharpe ratio and reduced drawdown by optimizing trading signals through **feature engineering**, **risk modeling**, and **backtesting**

PROJECTS

Visual Question Answering | *Python, PyTorch, ResNet50, Faster R-CNN, BERT, LSTM, Attention*

- Engineered multimodal **VQA models** by integrating **ResNet50**, **Faster R-CNN**, **BERT**, and **LSTM** for image-question understanding
- Applied **multi-head attention** and custom fusion layers to enhance image-text interaction, improving accuracy on **VQA benchmark datasets**

DailyResearcher.ai | *Next.js, Tailwind, Firebase, FastAPI, Shadcn UI*

- Built full-stack AI tool delivering personalized **arXiv** paper summaries with over **200 active users** onboarded
- Implemented plan-based access with **Firebase Auth** and **Firestore**, enabling user authentication, gated content, and interactive quizzes
- Launched responsive dashboard using **Next.js**, **Tailwind v4**, and **Shadcn UI**; increased user engagement through **gamified leaderboards**

TECHNICAL SKILLS

- **Languages:** Python, C/C++, JavaScript, Bash, HTML/CSS
- **Frameworks & Libraries:** FastAPI, Flask, React, Node.js, PyTorch, TensorFlow, NumPy, pandas, Matplotlib
- **Tools & Platforms:** Git, Docker, GitHub Actions, Nginx, Redis, DynamoDB, PostgreSQL, AWS (EC2, RDS, Lambda)

ACHIEVEMENTS

- **JEE Advanced (2022):** Secured All India Rank 3127 out of 1.5M+ candidates
- **JEE Mains (2022):** Achieved All India Rank 3600 (99.61 percentile among 10M+ candidates)
- **CBSE National Merit Certificate – Mathematics (2022):** Awarded to top 0.1 percentile of students nationwide
- **WorldQuant Alphathon – IIT Jodhpur (2023):** Ranked 14th in IIT Jodhpur
- **Amazon ML Summer School (2024):** Selected for Amazon's machine learning training cohort