

Shhaurya Jaiswal

+919369155131 | bt23csa021@iiitn.ac.in | LinkedIn: shhaurya-jaiswal | GitHub: Thebinary110

EDUCATION

Indian Institute of Information Technology (IIIT) Nagpur 2023 – 2027
B.Tech in Computer Science and Engineering (AI/ML Specialization) 3rd Year
– Relevant Coursework: Deep Learning, Cloud Computing, Distributed Systems, Containerization Technologies, ML Systems Design, Infrastructure Management

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, TypeScript, C/C++, Bash scripting, SQL
Machine Learning & NLP: scikit-learn, TensorFlow, PyTorch, pandas, numpy, Transformers, LLM optimization
Web Development: ReactJS, Node.js, Flask, REST APIs, Responsive UI Design
Cloud & DevOps: Docker, Kubernetes, AWS (S3, EC2, Lambda), Terraform, CI/CD pipelines, Git
Operating Systems: Linux (Ubuntu, Kali), Unix shell environments
Databases: PostgreSQL, MySQL, MongoDB, PineCone, ChromaDB
Development Skills: Back-End Development, Scalable Architecture Design, System Optimization, Performance Tuning
Data Analysis & Reporting: Microsoft Excel, Google Sheets, Trend Analysis, KPI Dashboards, Data Storytelling
Soft Skills: Analytical Thinking, Clear Communication, Team Collaboration, Attention to Detail, Resource Optimization

WORK EXPERIENCE

Ministry of Home Affairs, Government of India Dec 2024 – Feb 2025
AI Infrastructure Intern – Linguistic AI & ASR Systems Remote
– Designed and implemented containerized deployment pipeline for LLM-based ASR systems on Kubernetes
– Optimized Docker images and resource allocation for NLP models, reducing inference latency by 31%
– Developed monitoring systems for LLM performance across multiple languages using Python and Bash
– Automated resource scaling and container management using Kubernetes, improving system stability

MentorBox.ai Mar 2025
Full-Stack Intern – Web Platform Development Remote
– Built and deployed a full-stack website using **Python Flask**, enabling personalized mentorship matching
– Integrated **Docker** for development and deployment, ensuring consistent containerized environments
– Used **AWS Lambda** and **S3** for serverless backend processing and content storage
– Optimized RESTful APIs and database interactions to support dynamic content rendering and user analytics
– Implemented CI/CD pipelines for automated testing and production deployment

Research, IIIT Nagpur Jun 2024 – Aug 2024
ML Infrastructure Intern – Medical Imaging AI Nagpur, India
– Managed deployment infrastructure for deep learning models used in medical image analysis
– Optimized GPU resource allocation and scheduling, improving inference throughput by 24%
– Implemented containerized application workflows to enhance cross-team collaboration and model deployment

PROJECTS

LLM-DeploymentKit – Containerized LLM Deployment Framework
Docker, Kubernetes, Python, TensorFlow, PyTorch

- Built a scalable infrastructure system for deploying and managing multiple LLMs in production
- Implemented container optimization techniques to reduce image size by 47% and improve startup time
- Designed a monitoring and logging system for tracking model performance and resource utilization

AutoLegalBot – Containerized LLM Application Platform
Docker, Kubernetes, LangChain, Pinecone, Python

- Designed and deployed scalable infrastructure for LLM-based legal document processing
- Implemented efficient container management for multi-stage LLM pipeline with vector database
- Optimized resource utilization for batch processing and real-time inference workloads

CloudVault – Secure Infrastructure-as-Code Platform

Docker, Kubernetes, Terraform, AWS, PostgreSQL

- Built containerized application platform with infrastructure-as-code for automated deployment
- Implemented resource optimization strategies reducing cloud compute costs by 38%
- Designed monitoring and scaling infrastructure for managing application performance

ACHIEVEMENTS

Cleared IIT-JEE Main, ranked in the top 3% of candidates nationwide among 1.2 million applicants

2nd Runner-Up, IIT Guwahati Hackathon: Developed cloud-native containerized solutions in a 24-hour sprint

1st Place Winner, HackX @ IIIT Nagpur: Built a Kubernetes-powered ML infrastructure system

Infrastructure Optimization Challenge Winner: Reduced inference latency by 43% in containerized LLM deployment

EY Techathon Ideation Qualifier: Recognized for cloud-native infrastructure innovation

LEADERSHIP EXPERIENCE

Technical Community, IIIT Nagpur

2023 – Present

Cloud-Native Special Interest Group Lead

Nagpur, India

- Led workshops on containerization technologies and Kubernetes for ML/AI applications
- Collaborated with team members to develop best practices for cloud-native ML infrastructure