YUVRAJ SINGH RATHORE

+81 70-3815-6464 | yuvrajsinghrathore2118@gmail.com | github.com/yuvi2118 | Aizu-Wakamatsu, Japan

SUMMARY

M.S. Computer Science student at the University of Aizu, Japan (graduating Mar 2028; available as a 2028 new graduate), based in Japan. Software and security engineer: shipped Python/FastAPI services on PostgreSQL handling 50,000+ daily requests, and research AI-based multi-stage attack detection for industrial control systems (first-author paper under review at Information Fusion). Hands-on with Linux (RHCSA), Docker, Kubernetes, AWS and LLM guardrails; uses AI tools daily to build and measure faster. AWS Certified Developer.

EDUCATION

University of Aizu — M.S. Computer Science

Apr 2026 – Mar 2028 (Expected) | Aizu-Wakamatsu, Japan

- Research: multi-stage cyber-physical attack detection for industrial control systems using Transformer and diffusion-based models and information fusion (TEP and SWaT datasets).

Arya College of Engineering and IT, Jaipur — B.Tech, Computer Science

Oct 2021 – June 2025 | CGPA: 8.5

TECHNICAL SKILLS

Security and Detection: Multi-stage attack detection, cyber-physical and ICS security, anomaly detection, log and time-series analysis, AI system security (NeMo Guardrails, Guardrails AI), IAM

Backend and Data: Python, Bash, SQL, PostgreSQL, SQLAlchemy, FastAPI, Flask, REST APIs, Redis, Celery, Git, pytest, automation scripting

Linux, Cloud and DevOps: Linux administration (RHCSA), AWS (Lambda, EC2, EKS, IAM, Bedrock, SageMaker), Docker, Kubernetes, Terraform, GitHub Actions, CI/CD

AI and Machine Learning: PyTorch, Hugging Face Transformers, diffusion models, scikit-learn, LLM fine-tuning (PEFT, QLoRA), RAG, LangChain, LangGraph, MCP, Langfuse

Certifications: AWS Certified Developer Associate | Red Hat Certified System Administrator (RHCSA)

PUBLICATIONS

- **Rathore, Y.S.** and Saxena, N. — "CDIF: Cross-Domain Information Fusion Framework for Multi-Stage Cyber-Physical Attack Detection in ICS" — *Information Fusion (Elsevier)*, submitted 2026. [Under Review]

WORK EXPERIENCE

Divecho Technologies — Junior Python Backend Developer

Feb 2025 – Feb 2026 | Internship

- Built and operated RESTful APIs (Python + FastAPI) with SQLAlchemy and PostgreSQL handling 50,000+ daily requests; async task pipelines (Celery + Redis) cut API response time by ~35%.
- Worked in a team using Git, code review, GitHub Actions CI/CD and Docker; maintained greater than 85% pytest coverage across core services.
- **Stack:** Python, FastAPI, PostgreSQL, SQLAlchemy, Celery, Redis, Docker, GitHub Actions, pytest

PROJECTS

Domain-Adaptive LLM Fine-Tuning for Cybersecurity with PEFT and QLoRA

Personal Project | 2026

- Fine-tuned Mistral-7B on 12,000 cybersecurity Q and A samples with QLoRA (70% GPU memory reduction); DeepEval measured domain accuracy before and after; served with vLLM on AWS EC2 behind FastAPI.
- Added NeMo Guardrails to keep the model on topic and block out-of-scope requests. **Stack:** Python, PyTorch, Hugging Face PEFT, QLoRA, DeepEval, vLLM, NeMo Guardrails, AWS EC2

End-to-End Agentic RAG Pipeline with LangSmith and Langfuse Observability

Personal Project | 2025–2026

- Built a multi-stage agentic Q and A system (LangChain + LangGraph) with hybrid retrieval, a self-reflection loop and citation-grounded answers; deployed on AWS Lambda with a FastAPI service.
- Guardrails AI hallucination filtering cut off-topic responses by ~60%; Langfuse and LangSmith tracing for latency and cost; RAGAS evaluations run in CI. **Stack:** Python, LangChain, LangGraph, Guardrails AI, Pinecone, AWS Lambda

Industry-Grade MCP Client — Composable AI Tool Orchestration Platform

Personal Project | 2026

- Built a production MCP (Model Context Protocol) client in Python: JSON-RPC 2.0, stdio and SSE transports, typed schema validation, tool registry; Langfuse spans on every tool call for traceability.
- Docker Compose deployment with full pytest coverage. **Stack:** Python, Anthropic SDK, MCP SDK, FastAPI, Langfuse, Docker, pytest

Recall — Offline Android App for Interview Prep (Kotlin, Jetpack Compose)

Personal Project | 2026 | github.com/yuvi2118/recall

- Built an open-source native Android app (Kotlin, Jetpack Compose) that turns a 301-question GenAI and ML interview bank into an endless swipe feed with model answers; fully offline, with a section index and search, saved and review-later queues, and streaks.
- Wrote a Python pipeline that parses the source PDF into structured JSON bundled in the app; added daily reminder notifications and a release build tuned for smooth scrolling. **Stack:** Kotlin, Jetpack Compose, Gradle/R8, AlarmManager, Python