Maahin Rathinagiriswaran

New Brunswick, NJ | (732) 626-1051 | rmaahin2000@gmail.com | LinkedIn | GitHub | Portfolio | Google Scholar

Summary

AI/ML Engineer with 4+ years of production ML and software engineering experience, specializing in Agentic AI systems, RAG pipelines, and Applied ML. Experienced in building end-to-end systems spanning data curation, multi-agent orchestration, LLM fine-tuning, and production serving via vLLM.

Education

Rutgers University - New Brunswick — New Brunswick, NJ
Master of Science in Electrical and Computer Engineering — *May 2026*
- Specialization: Machine Learning
- Coursework: Computer Vision, Generative AI, Reinforcement Learning, Multimodal ML, 3D Imaging & Reconstruction

Princeton University — Princeton, NJ
Exchange Student - Advanced Computer Vision — *Jan. 2026*

PES University — Bangalore, India
Bachelor of Technology in Electronics and Communication Engineering — *May 2021*

Experience

WINLAB — New Brunswick, NJ
Machine Learning Intern — *May 2025 – Aug. 2025*
- Built a cross-modal gait re-identification system by training a Siamese Transformer for Video-IMU alignment and a gradient boosting based activity classifier, resulting in 0.98 AUC-ROC and 95% 3-class accuracy
- Engineered a relational **PostgreSQL** database and **Python** data pipeline to manage multimodal (IMU and Video) person re-identification data, resulting in a 90% reduction in post synchronization time across a dataset of 300+ videos and 21 subjects

PricewaterhouseCoopers (PwC) — Bangalore, India
Machine Learning Engineer — *Sept. 2021 – Aug. 2024*
- Built **KQL** pipelines to extract and clean 50,000+ Sentinel security alerts, creating a **Pandas** preprocessed dataset in **PostgreSQL**
- Trained an ensemble alert-triage classifier with **Scikit-Learn** (Random Forest + Gradient Boosting) on structured log features, reducing false-positive rates by 18% across 3 client SOC environments
- Fine-tuned **RoBERTa-base** using **PyTorch** and **HuggingFace** on 40,000+ raw alert descriptions to better capture severity, boosting F1 by 12% over the tabular-only model
- Containerized the end-to-end inference pipeline with **Docker** and deployed it as a **FastAPI** microservice on **Azure ML**, that cut average analyst triage time by 5 minutes per incident

Pi School of AI — Rome, Italy
AI Intern — *Sept. 2022 – Nov. 2022*
- Fine-tuned a Vision Transformer with **PyTorch** and **HuggingFace** for multilingual OCR, achieving 88% accuracy and 90% reduction in manual processing time; containerized with **Docker** and deployed via a **RESTful API** for production

Projects

SyncUp: Agentic University Project Management Tool | *Python, LangGraph, MCP, ReAct, LLM-as-a-Judge*
- Architected an agentic project management tool for university student group projects using **LangGraph**, **ReAct** agents, and **MCP** with 2 **LLM-as-a-Judge** evaluators, achieving 85% alignment with professor ratings
- Integrated NVIDIA **NeMo Guardrails** into the LLM orchestration layer to safely process external data from 4 integration sources (GitHub, Google Calendar, Google Docs, Trello), mitigating prompt injection attacks, sensitive data leakage, and jailbreak attempts

PaperFlow: Agentic Research Papers to Code Translator | *Python, LangGraph, vLLM, RAG, MCP*
- Architected a **LangGraph**+**MCP**-orchestrated pipeline to autonomously parse 600+ arXiv papers and benchmark generated code implementations against corresponding GitHub repositories, achieving 60% functional correctness on Papers With Code tasks
- Fine-tuned Qwen-2.5 14B via **LoRA** on 15,000+ paper-code pairs and deployed via **vLLM**, improving methodology-to-code alignment by 28% over baseline

SentinelGraph: Cyber Attack Path Analyzer | *LangGraph, Python, Neo4j, GraphRAG*
- Engineered a scalable Knowledge Graph pipeline using **Neo4j** to ingest, unifying 100,000+ data points into a unified threat topology
- Developed an agentic **RAG** system using **LangGraph** that interfaces with the graph database and LLMs, allowing analysts to query complex attack paths and reducing average threat investigation time by approximately 40%

Technical Skills

Programming Languages: Python, C, C++, SQL, MATLAB
ML Frameworks: PyTorch, TensorFlow, HuggingFace, Scikit-Learn, OpenCV, Pandas
AI Engineering: LangChain, LangGraph, LangSmith, RAG, GraphRAG, MCP, ChromaDB, vLLM, LLM-as-a-Judge
Tools: Docker, Microsoft Azure, FastAPI, PostgreSQL, Neo4j, Kubernetes, Git, ROS, LaTeX