

KHEN BO KAN

+82 10-8127-5897 hyonbokan@gmail.com Khen Bo Kan hyonbokan My Portfolio

Overview

Applied AI / LLM Engineer building production **LLM** and **agentic** systems. Work centers on autonomous **AI agent** pipelines, **RAG** retrieval, and **agent memory** across **OpenAI**, **Anthropic**, and **Google** models, with a focus on reliability, cost, and latency in production. Skilled in **Python**, **FastAPI**, **context engineering**, **fine-tuning** open-source LLMs, **LLM evaluation**, and **MCP** for external tool and data integration. Published in **IEEE Network Magazine** on fine-tuning LLMs for 5G network analysis.

Experience

AI Engineer — Nethermind – AuditAgent Project

Mar. 2025 – Present

Demerzel Solutions Limited

Remote (London, UK)

- Contributed to the development of **AuditAgent**, an AI-powered system for detecting vulnerabilities in smart contracts, trusted by leading protocols and institutions including **Uniswap**, **Lido**, **CMTA**, and **UBS**.
- Built an autonomous, multi-stage **LLM agent pipeline** for automated vulnerability detection and validation, reaching **40% recall** of professional audit findings and **67% on EVMBench**, scanning smart contract codebases of up to **12k lines** per run with checkpointing that resumes a scan after failure.
- Designed and implemented a **custom Chain-of-Thought (CoT) schema** to structure reasoning steps, enhancing model focus and vulnerability **detection performance by 22%**.
- Co-designed persistent **agent memory** and implemented its structuring and retrieval of audit context such as invariants, threat hypotheses, and prior findings across scans, through a unified **RAG** indexer that serves both code and memory, cutting prompt size by a **median 27%** while preserving deep code dependencies.
- Collaborated on the false positive reduction pipeline, achieving a **25% latency reduction** and **30% cost savings**.
- Built an in-house, provider-agnostic **LLM** layer reused across analysis workflows that routes across **OpenAI**, **Anthropic**, and **Google** models behind one interface, with structured outputs, a client-side tool-use loop, streaming, and automatic retry/fallback, without relying on a third-party SDK.

AI Network Lab Researcher — AI & LLMs for Network Analysis

Mar. 2023 – Feb. 2025

Chungnam National University

Daejeon, South Korea

- Developed an AI-powered web application for BGP anomaly detection, integrating a **few-shot prompted GPT** model and a **fine-tuned LLaMA** self-hosted with **vLLM** behind an **OpenAI-compatible API**, reducing manual analysis time by **96%** through automation. The application, built with a **React** frontend and **FastAPI** backend that streams responses over **SSE**, automates complex BGP analysis workflows, enabling users to detect anomalies through simple natural language queries.
- Published research in **IEEE Network Magazine** titled "*Mobile-LLaMA: Instruction Fine-Tuning Open-Source LLM for Network Analysis in 5G Networks*".
- Developed a dataset for **LLM training**, improving BGP analysis accuracy by **90 points** post-finetuning.
- Conducted **instruction fine-tuning** of open-source LLMs, significantly enhancing model adaptability and performance in network-related tasks.

Education

Chungnam National University

Mar. 2023 – Feb. 2025

Master of Science in Computer Science

Daejeon, South Korea

- Area of Research: Data Networks, LLM, Generative AI, Border Gateway Protocol (BGP), Federated Learning
- Academic Performance: Cumulative GPA: 4.25/4.5 (Percentage: 97.4%)

Technical Skills

- LLM & Agents:** AI Agents, RAG, Prompt Engineering, Fine-tuning, LLM Evaluation, LLMOps, Self-hosting (vLLM, TGI), MCP, PyTorch, Hugging Face
- Backend & Data:** Python, Java, FastAPI, Django, Spring Boot, PostgreSQL, MongoDB
- Frontend & DevOps:** TypeScript, React, Next.js, Docker

Language Skills

- English, proficient, IELTS 8.0
- Russian, native
- Mandarin Chinese, proficient, HSK 6
- Korean, fluent, TOPIK 4