

Minah Kim

010-9172-0584 | mg980504@gmail.com | GitHub | Tech Blog | Portfolio | LinkedIn

EXPERIENCE

TagHive | AI Engineer Jul 2023 - Present

An educational solution company targeting students from grade 3 to 10 in India

- Implemented and released a beta version of a data analysis agent tailored to user requirements.
- Unlike LLMs performing single tasks, the agent improved accuracy by designing clear tasks and contextual prompts.
- Built a precise analysis process by using custom tool calling for specific internal calculations.
- Managed learning data using Cronjob and AWS S3; detailed analysis for data within 6 months, and summary statistics for data over 1 year to optimize analysis speed and data handling.
- Developed a validation system for a fine-tuned LLM (Gemma2 9B)
- Automated validation with a Judge LLM, reducing human resource usage by 66.7%.
- Evaluated 4 models by comparing Reference answers and model responses using LLM-as-a-Judge paper.
- Achieved 81% agreement between Judge LLM and human evaluators, proving evaluation reliability.
- Improved LLM fine-tuning and RAG pipeline
- Reduced API cost by 30% using OpenAI model for parsing textbook PDFs with a batch architecture.
- Designed optimized RAG pipeline and VectorDB suitable for datasets.
- Repeatedly improved performance using Ragas evaluation system.

- Served fine-tuned Llama3 8B model in Azure (multi-GPU) with DeepSpeed memory optimization.
- Developed and researched Deep Knowledge Tracing (DKT) model
- Implemented a student answer time prediction model using a custom DKT model combining Transformer and LSTM.
RMSE improved by 66.46%, and learning speed by 88.24%.
- Modified model structure to pass probabilities without sigmoid activation, enabling multitask learning.
- Third author of "Multi-task Transformer with LSTM Model for Question Set Generation" paper and applied for two patents.

PROJECTS

Resume-based Job Recommendation and Suggestion Service Oct 2024 - Present

Fake Research Lab, 9th-10th cohorts

- Led the 5-person team and project planning; built a collaborative environment with GitHub Discussion and Notion Sprint.
- Used LangChain Semantic Chunker for semantic chunking of job descriptions; original JD stored in DB and chunks vectorized in ChromaDB, improving original page matching.
- Generated a matching report between resume and JD, providing 3-line summaries and highlights for better readability.
- Analyzed requirements and preferences to generate scored feedback so applicants can easily identify strengths and gaps.

Turtle Neck Diagnosis and Stretching Service using AI Dec 2022 (5 weeks)

Elis AI Track, 5th cohort

- Captured ~1300 webcam images via Zoom; used Movenet for pose estimation and extracted upper body data.
- Trained model using XGBoost classifier with SMOTE and K-Fold cross-validation (F1 score: 0.92).
- Built FastAPI server and implemented real-time client-server prediction with Socket.IO.

Object Detection and Distance Estimation in Farm Environments Jul - Dec 2021 (6 months)

Chonnam National University

- Collected and labeled farm environment data; trained YOLOv5 model (F1 score: 0.934).
- Applied cutout augmentation for a 0.04 mAP increase.
- Used center 1/9 region of bounding box for depth data to reduce distance error (RMSE) from 0.22 to 0.04.
- Successfully tested object detection and distance estimation in new farm environments and published paper.

CERTIFICATIONS

- SQLD (SQL Developer Certification) Sep 2023
- ADsP (Data Analysis Semi-Professional Certification) Feb 2022
- TOEIC 880 Sep 2024
- OPIc IH Mar 2025

EDUCATION

- Elis AI Track, 5th cohort Jun - Dec 2022 (6 months)
- Chonnam National University, College of Agricultural and Life Sciences, Major in Bio-Industrial

Machinery Engineering Mar 2017 - Feb 2022 (4 years)

OTHER EXPERIENCE

- CMES | R&D Vision Algorithm Engineer Intern Apr - May 2023 (2 months)
- Kimdaejung Convention Center | Innovation Project Team Intern Mar - May 2022 (3 months)
- IndieJ | AI Music Recommendation Service, Marketing Intern Sep - Dec 2020 (4 months)