Oyeon Kwon

☐ +82 010 4024 5879 | @ data0311@gmail.com | 🔗 LinkedIn | ⌂ GitHub | 🌐 Portfolio | 📍 South Korea

EDUCATION

Carnegie Mellon University   Pittsburgh, USA
M.Eng. in Biomedical Engineering;   *Sep 2024 – Present*

Korea University   Seoul, South Korea
M.Eng. in Brain and Cognitive Engineering;   *Graduation Date: 2018*
Advised by Klaus-Robert Muller, Cuntai Guan, and Seong-Whan Lee*(Major)   *GPA: 4.44/4.50*

SKILLS

Programming Languages: Python, TypeScript, MATLAB, Rust, CUDA, C/C++
Frameworks & Libraries: FastAPI, PyTorch, TensorFlow, Triton, JAX, Django, Next.js, Nest.js, ONNX, PEFT, DeepSpeed, Langchain, LlamaIndex
Databases: OpenSearch, MySQL, MongoDB, VectorDB (Pinecone, Qdrant, Milvus, Weaviate)
MLOps & Cloud Services: Git, Docker, BigQuery, AirFlow, Kubeflow, k8s, MLFlow, Kafka, AWS, GCP, Azure
AI & Machine Learning: 1D/2D Cls./Pred., 3D anim. generation., LLM Finetuning, RAG Flow Engineering
Research Achievements: Google Scholar: 900+ citations, h-index 7, i10-index 7, US Patent +3, KR Patent +5

EXPERIENCE

FluentT   Seoul, South Korea
Product Lead & AI Engineer   *Sep 2023 – Present*
- Led 3D Realistic & Conversational Avatar Projects: http://talkmotion.ai (Unity, Development, Design, AI)
- Set up scalable and high-availability initial cloud server architecture and management (Terraform)
- Trained audio-driven diffusion-based 3D animation generation model, reducing inference time from 10s to 0.1s
- Deployed LLM - TTS - 3D animation model in clouds using Triton Ensemble for real-time service
- Developed real-time chatbot agent using RAG-flow engineering (Adaptive RAG) and layered indexing/search
- Designed automatic data acquisition and response evaluation pipelines
- Built MLOps pipeline level 2 (Basic CI/CD Process)

Scrambler   Seoul, South Korea
Co-Founder, CTO/CPO   *July 2022 – July 2023*
- Launched 1:1 local Korean matching service achieving 4x month-on-month growth from zero-to-one
- Secured pre-seed funding from Antler, leading to $0.1M in funding
- Developed cross-platform app (Flutter & Firebase) with integrated video calls (Agora)
- Improved 50% retention rate by analyzing customer behavior with UXCam and optimizing initial content lock-in
- Achieved CPA of $0.15 (90% lower market average) by tracking 20-30 female users and rescheduling SNS content

VUNO   Seoul, South Korea
Part Lead & AI Researcher   *June 2018 – April 2022*
- Participated as the main AI researcher from Series B funding to IPO process
- Developed an AI early warning system for cardiac arrest, saving lives by early detection
- Led E2E AI production process from study design, data processing, feature engineering, modeling, and evaluation
- Built an RL-based AutoML system demonstrating performance comparable to expert-trained models
- Implemented production-level optimizations; batch sampling (0.01% positive samples), validation set creation, etc.

Korea University   Seoul, South Korea
Researcher   *June 2016 – Aug 2018*
- Published in a top-tier CS journal (1/104) as the first author (Reduced motor imagery calibration time in BCI)
- Contributed to open source ('PatternRecognition/OpenBMI') and open dataset ('gigadb.org/dataset/100542')