

# EESUN MOON

(646) 476-0342 | [eesun.moon@columbia.edu](mailto:eesun.moon@columbia.edu) | [LinkedIn](#) | [GitHub](#) | NY, United States

## EDUCATION

### Columbia University

MS in Computer Science, GPA: 3.92/4.0

New York, NY

Expected Dec 2025

- Courses: Applied Machine Learning, Natural Language Processing, Spoken Language Processing, Computer Vision, Database

### Sejong University

BS in Intelligent Mechatronics Engineering, BE in Data Science, GPA: 4.4/4.5

Seoul, South Korea

Feb 2024

- Courses: Artificial Intelligence, Computer Networks, Operating Systems, Image Processing, Data Structures, Web Programming
- Teaching Assistant: Algorithms using C programming, Python fundamentals

## TECHNICAL SKILLS

### Programming & Databases

Python, C, R, Java | MySQL, PostgreSQL, MongoDB

### Machine Learning Frameworks

TensorFlow, Keras, PyTorch, Scikit-Learn, Hugging Face, OpenAI, LangChain

### Development & Cloud

Git, Docker, Linux (Ubuntu), FastAPI, Flask, Cloud Deployment (AWS, GCP)

### Hardware & Data Tools

GPU, NPU, ONNX | Pandas, NumPy, Selenium, Matplotlib

## PUBLICATIONS

[1] Eesun Moon, A.S.M Sharifuzzaman Sugar, Hyung Seok Kim, "Multimodal Daily-life Emotional Recognition Using Heart Rate and Speech Data from Wearables," *IEEE Access*, vol. 12, pp. 96635-96648, 2024. [DOI](#)

[2] Taein Kim, Eesun Moon, Hyeon Kang, Hyung Seok Kim, "OMER-NPU: On-device Multimodal Emotion Recognition on Neural Processing Unit for Low Latency and Power Consumption," *Neural Computing and Applications* (in press).

## PROFESSIONAL EXPERIENCE

### Samsung Research America

AI Algorithm/NPU Simulator Research Scientist

Mountain View, CA

Jun 2025 – Expected Aug 2025

### Humaner: Human-centered AI Software Development [\[GitHub\]](#)

Seoul, South Korea

Machine Learning Engineer

Mar 2024 – May 2024

- Built and deployed Q&A-based support message generator for soccer players using OpenAI and LangChain on Dockerized AWS EC2, enabling real-time interaction with 500+ live users and increasing **satisfaction by 20%**
- Tuned prompts based on post-deployment survey feedback to improve message relevance and personalization

### Sejong University, Mobile Intelligent Embedded System Laboratory [\[GitHub\]](#)

Seoul, South Korea

Research Assistant

Sep 2021 – Mar 2024

- Led multimodal emotion recognition project for on-device AI using TensorFlow and MongoDB on Linux for government initiatives
- Optimized Keras-based deep models with score-based fusion of multimodal signals (heart rate, EEG, speech, image), achieving **99.68% classification accuracy** without increasing network complexity
- Deployed ONNX models on MLA100 NPU, reducing **power consumption by 3.12x** and **latency by 1.48x** for edge deployment
- Published papers in **IEEE** (Institute of Electrical and Electronics Engineers) [1] and **NCAA** (Neural Computing and Applications) [2] and demonstrated live deployment at **KIST** (Korea Institute of Science and Technology)

## PROJECTS

### Sentence Embedding Analysis in LLMs [\[GitHub\]](#)

Jan 2025 – May 2025

- Analyzed embedding interpretability in LLMs through Zipf-like cluster distributions, demonstrating intermediate layers form more structured groupings, with slopes steepening from -0.87 to -1.42 as domain specificity increased
- Evaluated clustering reliability using Hugging Face Transformers, PyTorch, and scikit-learn, finding intermediate-layer and domain-specific embeddings yield more coherent clusters

### CS Advising Assistant Chatbot with LLM, RAG, and Agentic Flow [\[GitHub\]](#)

Jan 2025 – May 2025

- Developed chatbot with local inference via DeepSeek on Ollama to **eliminate LLM API costs** and deployed on GCP for production
- Optimized RAG pipeline with LangChain and MCP server, integrating Agentic Flow for multi-step retrieval and tool-based reasoning

### Ranking-Based Spam Filtering on Social Networking Services [\[GitHub\]](#)

Mar 2022 – Jun 2022

- Spearheaded project to prioritize organic user posts over likely ads from social media, earning **1st place** in graduation competition
- Automated data collection with Selenium and implemented unsupervised clustering with cosine similarity-based ranking, achieving **0.8 intra-cluster similarity** as coherence indicator