

## RESEARCH INTERESTS

---

**Clinical NLP, Biomedical NLP, Multi-modal LLMs, XAI**

## EDUCATION

---

<b>Korea University</b> M.S. - Computer Science & Engineering; GPA: 3.89/4.0	Seoul, Republic of Korea [2023.03 - 2025.02]
<b>University of Lausanne</b> Exchange student in School of Biology	Lausanne, Switzerland [2018.02 - 2018.07]
<b>Korea University</b> B.S. - Biotechnology; GPA: 3.37/4.0	Seoul, Republic of Korea [2016.03 - 2022.02]

## ACADEMIC EXPERIENCE

---

<b>Boston Children's Hospital (Harvard Medical School)</b> Visiting Student (Advisor: Dr. Timothy A. Miller) <ul style="list-style-type: none"><li>Awarded scholarship from Korea University - KIAT Scholarship Program</li><li>Led a project on ICU patient mortality prediction using multi-modal LLMs</li></ul>	Boston, Massachusetts, USA [2024.10 - ongoing]
<b>Korea University</b> Research Intern (Advisor: Dr. Jaewoo Kang) <ul style="list-style-type: none"><li>Contributed to a project that using reinforcement learning for chest X-ray summary generation</li></ul>	Seoul, Republic of Korea [2022.12 - 2023.02]
<b>Korea University</b> Research Intern (Advisor: Dr. Chungho Kim) <ul style="list-style-type: none"><li>Contributed to a project on integrin activation regulation mechanisms in angiogenesis</li></ul>	Seoul, Republic of Korea [2021.02 - 2021.02]
<b>Korea Institute of Science and Technology (KIST)</b> Research Intern (Advisor: Dr. Hyunkyung Kim) <ul style="list-style-type: none"><li>Contributed to a project on epigenetic markers related to eye cancer</li></ul>	Seoul, Republic of Korea [2019.01 - 2019.02]

## INDUSTRY EXPERIENCE

---

<b>Kolon Lifescience</b> Clinical Trial Project Manager <ul style="list-style-type: none"><li>Managed long-term follow-up studies for clinical trials and the writing of clinical study reports</li><li>Coordinated with CROs, vendors, and clinical sites to meet timelines and compliance</li></ul>	Seoul, Republic of Korea [2022.01 - 2022.10]
<b>BIO Division at CJ CheilJedang</b> Research Intern in Purification Group <ul style="list-style-type: none"><li>Investigated amino acid purification processes (L-Arginine) using membrane filtration technologies</li></ul>	Seoul, Republic of Korea [2022.11 - 2022.12]
<b>Samsung Bioepis</b> Research Intern in Purification Group <ul style="list-style-type: none"><li>Conducted research on biosimilar purification processes using Protein A Chromatography</li></ul>	Incheon, Republic of Korea [2022.07 - 2022.08]

## HONORS AND AWARDS

---

- 1st Prize at ACL 2023 BioNLP Workshop**
  - Developed pre-trained vision-language model for chest x-ray radiology report summarization
- 1st Prize at HackSeoul 2024 Social Responsibility Track**
  - AI-Powered express bus reservation system for the digitally alienated elderly
- 3rd Prize at BioCreative VIII Track 3**
  - Extracted and normalized genetic phenotypes from dysmorphology physical examination entries
- 4th Prize at NACCL 2024 ClinicalNLP Workshop**
  - Developed reliable Text-to-SQL generation models for extracting information from EHR
- Top 9 at HackGlobal Grand Finals**
  - AI-powered platform facilitating social gatherings and community engagement for hotel guests

- [R] Yoon, W., Ren, B., Thomas, S., **Kim, C.**, Savova, G., Hall, M., G., and Miller, T. **Aspect-Oriented Summarization for Psychiatric Short-Term Readmission Prediction** *arXiv preprint*
- [R] **Kim, C.\***, Kim, H.\*, Park, S., Lee, J., Sung, M., and Kang, J. **Learning from Negative Samples in Generative Biomedical Entity Linking.** *arXiv preprint*, (\* Equal Contribution)
- [C] Kim, H.\*, **Kim, C.\***, Lee, H., Jang, K., Lee, J., Lee, K., Kim, G., and Kang, J. **KU-DMIS at EHRSQL 2024 : Generating SQL query via question templatization in EHR.** *NAACL 2024 (6th Clinical NLP Workshop)* (\* Equal Contribution)
- [C] Kim, H., **Kim, C.**, Sohn, J., Beck, T., Rei, M., Kim, S., Simpson, T. I., Posma, J. M., Lain, A., Sung, M., and Kang, J. **KU AIGEN ICL EDI@BC8 Track 3: Advancing Phenotype Named Entity Recognition and Normalization for Dymorphology Physical Examination Reports.** *Proceedings of the BioCreative VIII Challenge and Workshop 2023*
- [C] Lee, S., Jang, G., **Kim, C.**, Park, S., Yoo, K., Kim, J., Kim, S., and Kang, J. **Enhancing Clinical Outcome Predictions through Auxiliary Loss and Sentence-Level Self-Attention.** *IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2023*
- [C] Kim, G., Kim, H., Ji, L., Bae, S., **Kim, C.**, Sung, M., Kim, H., Yan, K., Chang, E., and Kang, J. **KU-DMIS-MSRA at RadSum23: Pre-trained Vision-Language Model for Radiology Report Summarization.** *ACL 2023 (22nd Biomedical NLP Workshop)*

---

## PROJECTS

### ANGEL Framework

- Codebase for the paper "*Learning from Negative Samples in Generative Biomedical Entity Linking*"
- Published the source code on GitHub and model checkpoints on Huggingface for public use
- Designed a biomedical entity linking model using a generative model with preference optimization

### Human-in-the-loop Large Language Model for Biomedical Hypothesis Generation

- Led a lab project exploring advanced methods for biomedical hypothesis generation
- Investigated the use of agent-based systems and RAG-based approaches to develop human-in-the-loop LLMs for biomedical hypothesis generation

### AI Platform for Precision Medicine in Diabetes and Its Complications Using Big Data

- Served as Sub-Leader in a laboratory project focused on precision medicine for diabetes
- Fine-tuned LLMs to deliver domain-specific, expert-level explanations by incorporating advanced medical knowledge, clinical guidelines, and patient-specific information for diabetes care

### Open-source LLM for Explaining Korean Cultural Heritage

- Led a lab project developing a multi-turn chatbot using Korean LLMs for personalized storytelling
- Pretrained multi-lingual LLMs and integrated RAG to improve contextual understanding
- Fine-tuned LLMs with audience preference graphs transformed into natural language queries for generating personalized, interest-based stories

---

## TEACHING

### Teaching Assistant

[2023]

- Introduction to data science course
- Machine learning course
- Supported in course instruction, grading, and mentoring students

---

## SKILLS

**Programming:** Python    **ML/DL Frameworks:** TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy  
**LLM & NLP:** Huggingface, OpenAI API, LangChain, LlamaIndex, RAG, vLLM    **Cloud:** Azure