Dongsuk Jang

Interdisciplinary Program in Bioengineering,
Graduate School of Engineering, Seoul National University
Convergence Bldg 323, 103 Daehak-ro, Jongno-gu

Email: jamesjang26@snu.ac.kr
Mobile: +82-10-5105-3184
LinkedIn: www.linkedin.com/in/jamesjang26
GitHub: github.com/JamesJang26

Education

- **Seoul National University** — Seoul, Korea
  M.S. Interdisciplinary Program in Bioengineering — *Mar 2023 - Feb 2025(Expected)*
  Research Interests: *Natural Language Processing, Multimodal AI, Medical AI*

- **Sungkyunkwan University** — Suwon, Korea
  B.S. Integrative Biotechnology — *Mar 2015 - Feb 2023*
  Micro Degree *Core-Bio A-School Track*

- **Haeundae Highschool** — Busan, Korea
  Established CBC(Chemical Biology Club), 3 years in Mathematical Essay Club — *Mar 2012 - Feb 2015*

Publications

- **Dongsuk Jang†, Hyeryun Park†, Jiye Son, Hyeonuk Hwang, Su-jin Kim, Jinwook Choi**: "Automated Information Extraction from Thyroid Operation Narrative: A Comparative Study of GPT-4 and fine-tuned KoELECTRA" In: American Medical Informatics Association (AMIA), Informatics Summit. 2024.

Projects

- **Table to Text Generation task research (Dec 2023 - Present)**: Table to Text generation research comparing Seq2Seq models with Hierarchical Encoder and diverse LLM finetuning methods(PEFT) such as QLoRA, etc.
- **Natural Language Search System on Clinical Data Warehouse (Feb 2024 - Present)**: Development of a search system that answers doctor's natural language questions about patient info. Converting medical natural language questions into best sql queries.
- **Automatic Generation of Thyroid Operation Records on Web(Mar 2023 - Feb 2024)**: Development of a web-based system (voice recognition using CLOVA Note, extraction of clinical info with KoELECTRA, normalization of extracted results, image generation) and comparison with GPT-4 few-shot setting.
- **CNN Image Classification Project(Mar 2023 - May 2023)**: Building basic CNN Image classification model that classifies sports balls vs human faces.
- **Word2Vec from Scratch(May 2023 - Aug 2023)**: Study the principle of word2vec and build up from scratch. Train the own model with the book 'Alice in Wonderland' and compare performance with original word2vec library.
- **Winter Internship Project(Dec 2022 - Feb 2023)**: Compare the performace of various models(m-LSTM as baseline, BERT based models such as BioELECTRA, DeBERTa, PubMedBERT, RoBERTa etc.) MedNLI Task with MIMIC-III Dataset and find out the best parameter settings.

Experience

- **Student Intern, Samsung Advanced Institute for Health Sciences and Technology**:
  Mar 2021 - Dec 2022
  - **About**: I started studying computer science subjects for the first time, realizing that much research in biology and medicine is also happening through the use of computers, which sparked a great curiosity and had the biggest influence on my current career choice.
  - **Supervisor**: Seonguk Seo(03/2021-05/2022), Wonchul Cha(06/2022-12/2022)
- **Research Intern, Liflex Science Inc.**:
  Jun 2022 - Aug 2022
  - **About**: With a great interest in entrepreneurship and business, worked as an intern at a startup company, learning about the overall practical aspects of a startup, especially in investment acquisition and R&D. Realized that more important than novel ideas is to concretize them and develop them into a business.
  - **Supervisor**: Jinhong Kim, Dongsoon Ahn
- **Military Service, Korean Augmentation to the United States Army, USAG-Daegu**:
  Jan 2018 - Sep 2019
  - **About**: Worked alongside the Korean and American military, understanding and coordinating the differences between each military and culture. Additionally, learned how to live a proactive life based on strong physical fitness and a regulated lifestyle.
  - **Role**: Human Resources Specialist(42A) & Senior KATUSA

Honors and Awards

- Next Generation BioHealth Leader Award at Sungkyunkwan University - Feb, 2023
- Best Intern Award at Samsung Advanced Institute for Health Sciences and Technology - August, 2021
- Commandants list at KATUSA training academy - April, 2018

Teaching Assistant

- **Computer Assisted Medical Data Processing in the Department of Medicine** 2024
  Python Basics, Data Analysis and Machine Learning(breast cancer dataset)
- **Selective Course for undergraduate students in the Department of Medicine** 2023
  Python Basics, Data Analysis using pandas(breast cancer dataset), ANN classification(breast cancer dataset)

Professional Services

- Reviewer for the AMIA 2024 Annual Symposium

Skills Summary

- **Languages**: Python, R, Matlab
- **Frameworks**: Pytorch, TensorFlow
- **Tools**: Git, Docker
- **Platforms**: Linux(Ubuntu), Windows
- **Soft Skills**: Leadership, Writing, Time Management