

## Internships

### Process-Aware Data Analytics Lab Jan 2024 to Present

Undergraduate researcher, in Kwangwoon University

- Data Science research based on Process mining and Text mining

## Project Experience

### Embedding Process Models Using Language Models Feb 2024 to Present

Collaborative Research, in PADA Lab

- Methodology for **Sbert Model** Fine-Tuning for Efficient Embedding Process Models & Process of Converting Tree Models to Text
- Similarity Analysis of Process Models Based on Semantic and Structural Differences
- Visualization of Embedded Results through Dimensionality Reduction
- Currently Writing a Paper Targeting the EAAI Journal

### Fashion Trend Analysis & Development of Image Search Service Feb 2024 to Oct 2024

Team Member, for *Capstone design* project for graduation

- Crawling data from three fashion websites & Automating Crawling Using **Apache Airflow**
- Developing a website for comparing and analyzing collected data
- Developing image search API using **Grounding Dino** & **SAM** models

### Development of a Language Model for Risk Analysis in Cybercrime Prevention Mar 2024 to Jun 2024

Team Leader, for Major Courses *Text Mining* project

- Using Crime Glossary Data for Terms Such as Drugs and Sexual Offenses.
- Fine-Tuning **BERT Classification Model** for Risk Category and Severity Analysis of Internet Post
- Providing Crime Prevention Solutions Including Relevant Legal Provisions and Penalties Based on Analytical Results to Authors
- Creating Chatbot for Counseling and Advice Using **GPT API** and **RAG** Technology

### Prediction of Thyroid Disorders via Patient Biometric Signals with Machine Learning Sep 2023 to Dec 2023

Team Member, for Major Courses *Data Mining* project

- Collecting Data Including Patient Test Information and Thyroid Symptoms
- Handling Variables Through Exploratory Data Analysis (EDA)
- Comparing Model Performance Based on Various Feature Selection Methods(**Stepwise**, **backward**, etc.) and Machine Learning Techniques(**Naive Bayes**, **XGBoost**, etc.)

Team Member, for Major Courses *Big Data Processing and Applications* project

- Constructing an Item Matrix Using Tags and Genre Information from MovieLens Data
- Building a System Based on Various Similarity Measures such as **TF-IDF**, **Pearson Similarity**, and **Jaccard Similarity**
- Implementing Big Data Parallel Processing Using **PySpark**

Team Member, for Major Courses *Data Mining* project

- Utilizing Various **Machine Learning Techniques** for Prediction, including Linear Regression, Decision Trees, and Random Forests
- Enhancing Model Performance through **Ensemble Methods**
- Interpreting Results Using **SHAP** and **LIME** for **Explainable AI (XAI)**

EDUCATION

B.S. in Information Convergence, Major in Data Science  
- Total GPA of 4.0 / 4.5, Major GPA of 4.14 / 4.5

Awards & Honors

Awards

2023.11 Dean’s List, for Academic Excellence

Kwangwoon Univ.

Honors

2024 Semester High Honors, Awarded to students with high achievements throughout the semester

2023 Semester High Honors, Awarded to students with high achievements throughout the semester

2023 Semester High Honors, Awarded to students with high achievements throughout the semester

SKILLS

Programming Language

Python

SQL

Deep Learning / Machine Learning

PyTorch

Pandas

Scikit-learn

Keras

PySpark

TensorFlow

Numpy

Matplotlib

Seaborn

Tools

- Visual Studio Code
- Google Colab
- Anaconda
- Jupyter
- Apache-Airflow
- Wandb

Cloud

- GCP
- AWS

Used at least once

- HTML/CSS
- Javascript