

# Byungjun Yoon

☎ +82 010 3309 3695 | @ bjyoon513@gmail.com | @ junyoon@postech.ac.kr | 🔗 LinkedIn | 🐙 GitHub | 📍 Seoul, Korea

## RESEARCH INTEREST

My research interest lies in various topics in **Machine Learning, Large Language Models**, and **their practical applications**. I am particularly interested in ML/AI topics that can enhance human productivity. Currently, the large language model and its factual content generation are my main focuses.

## EDUCATION

**Pohang University of Science and Technology**

*B.S. in Computer Science*

Pohang, Korea

*Feb 2019 – Present*

**McLean High School**

Virginia, USA

*Sep 2015 – Jun 2018*

## WORK

**Devsisters**

Seoul, Korea

*Applied Data Scientist, Software Engineer*

*Sep 2021 – Oct 2023, Full-time*

- Led the development of a Customer ROAS Prediction System using Deep Learning for data-oriented UA marketing decisions.
- Improved the legacy MCMC-based Life Time Value Model.
- Implemented and managed a PPO-based puzzle game automation AI.
- Developed a log-based user churn prediction algorithm for Cookie Run: Kingdom.
- Developed a distributed PyTorch dataloader wrapper for Parquet datasets in a distributed data environment.
- Developed a front-end webpage with Next.js.

**Naver Webtoon**

Pangyo, Korea

*Research Intern, W-Tech Research Team: AI Automation*

*Jun 2021 – Aug 2021, Internship*

- Researched unsupervised clustering of Webtoon drawing styles related to deformation to provide classification of art styles for ML tasks.
- Investigated Webtoon drawing style embedding algorithms and similarity metrics based on metric learning.

**Reziena**

Seoul, Korea

*Intern*

*Jun 2020 – Aug 2020, Internship*

- Researched the backend and ETL pipeline for data from IoT devices, developed a Skin Type Classification Algorithm, analyzed AWS Infra security, and implemented a Firebase authentication system.

## RESEARCH

**DARPA AI Cyber Challenge**

Team Atlanta

*Lead by Taesoo Kim @GaTech*

*Sep 2023 – Present, ML Task Member*

- Working as a team member in Team Atlanta consist of SSLab@GaTech, Hacking Lab@KAIST, ML Lab@POSTECH, Samsung Research, and others
- Researching ML pipeline of automatic CWE patch (including CWE detection, patch generation)

**Enhancing retrieval-augmented generation by prompting model: A tri-model framework for factual language generation**

DI Lab

*Advisor: Hwanjo Yu*

*Sep 2023 – Present, Internship*

- Researching a tri-model framework, which inserts a prompting model between a retriever model and a generator model to improve the RAG framework in a black-box frozen generator environment.

**Lung Abnormalities Detection**

AIMI Lab

*Advisor: Wonhwa Kim*

*Oct 2020 – Jun 2021, Internship*

- Researched lung abnormality detection in chest x-ray images using R50-FPN for localization and abnormality classification.

**Undergraduate Research Participation**

CV Lab

*Advisor: Suha Kwak*

*Dec 2020 – Feb 2021, Internship*

- Studied Metric Learning, Computer Vision, and CS231n.
- Reproduced related papers: [1], [2].

## SKILLS

---

**Technologies:** PyTorch, PyMC, Pyro, Slurm, TensorFlow, PySpark, Git, Docker, Kubernetes, Helm, Next.js, React, Unity ML-Agent, Flask, Node.js

**Languages:** Python, Java, C/C++, JavaScript/TypeScript, Solidity, Rust

## AWARDS & ACHIEVEMENTS

---

**1st Place Award in the Postech Open Innovation Big Data Challenge**

## PROJECTS

---

### **Deep Learning (CSED538) In-class Image Classification Competition | *Github***

- Developed an image classification framework based on ViT and RandAugment
- Achieved first place in the competition

### **Personality Prediction AI and App Development | *GitHub***

- Created a computer vision project to classify a person's MBTI with vision models
- Deployed a web application to serve our model in a web environment.

### **Solar Energy Prediction System | *Github***

- Developed a multivariate time-series prediction model for predicting solar energy production
- Received the First Place Award in the Postech Open Innovation Big Data Challenge

## RELEVANT COURSEWORK

---

**Major coursework:** Deep Learning (CSED538), Deep Learning NLP (CSED554), Implem. & Accel. for Machine Learning (CSED510), Software Design Methods (CSED332), Artificial Intelligence (CSED342), Data Analysis Using Tools (CSED490A), Operating Systems (CSED312), Computer Architecture (CSED311), Automata & Formal Languages (CSED341), Algorithm (CSED331), Object Oriented Programming (CSED232), Data Structure (CSED233), Discrete Math (CSED261), Digital Sys. Design (CSED273), Intro. to Computer SW Systems (CSED211), Applied Linear Algebra (MATH203),

## MENTORING

---

POSTECH Intro. to Artificial Intelligence (AIGS101) TA

2nd POSTECH ON-LINE Hacking & Cyber Security Camp Mentor

5th POSTECH Computational Thinking Camp Mentor

## ORGANIZATIONS

---

**PLUS (POSTECH Labatory for Unix Security)**

*Member*

*Aug 2019 – Present*

**PDAO (POSTECH Decentralized Autonomous Organization)**

*Committee Member*

*Apr 2022 – Present*