



# SaeChan Oh

(+82) 10-8731-1883 | dhocks@naver.com | github.com/newfull5

## EDUCATION

---

### Jeju National University

Jeju, Korea

B.S in Computer Engineering (Mar 2018 - Jul 2024)

- *Coursework:* programming with Python, Java, Linear Algebra, Data Structure, Operating System, Computer Network, Computer Graphics, Computer Architecture, Machine Learning

### Machine Learning Lab, Jeju University

Jeju, Korea

Undergraduate research student (Mar 2020 - Nov 2020)

### West Visayas State University

Iloilo, Philippines

Language training abroad (Dec 2020 - Jan 2021)

## WORK EXPERIENCE

---

### NCSOFT Corp.

Seongnam, Korea

Internship as NLP Engineer (Sep 2024 - )

- Compose Korean-Vietnam parallel data.
- Compose RAG data.

### TUNiB Inc.

Seoul, Korea

Internship as NLP Engineer (Jun 2021 - Dec 2021)

- Compose Korean large training data (Polyglot-ko, TUNiB-Electra).
- Detecting Korean hate-speech and severity

## Nanoom Energy Inc.

Jeju, Korea

Internship as Software Engineer (Jun 2020 - Jan 2021)

- Prediction for electrical energy of solar panel

## Aurora House Inc

Jeju, Korea

Internship as Software Engineer (Mar 2019 - Jul 2019)

- Developing an Application for promoting Jeju support project

## ACTIVITY

---

### JipHyeonJeon

NLP paper review meeting (May 2021 - Dec 2020)

- had a presentation about BERT paper
- [github.com/jiphyeonjeon/season1](https://github.com/jiphyeonjeon/season1)

### Geultto season5

Technical article posting and peer review meeting (Oct 2020 - May 2021)

- had a presentation about BERT paper
- [github.com/jiphyeonjeon/season1](https://github.com/jiphyeonjeon/season1)

### UMLC2020 x BlockChain Camp

Jeju, Korea

Inter university student machine learning & blockchain camp (Aug 2020)

- Administor

## OPEN SOURCE PROJECT

---

### Transformers

Open-source contribute HuggingFace (Sep 2024)

- write Korean tutorial document

### NLLB-200 Distilled 350M en-ko

English to Korean translation model based on facebook NLLB-200 (Apr 2024)

- knowledge distillation
- [github.com/newfull5/NLLB-200-Distilled-350M-en-ko](https://github.com/newfull5/NLLB-200-Distilled-350M-en-ko)

## TUNiB-Electra

Pre-trained Korean encoder model (Sep 2021)

- compose training dataset and preprocess
- [github.com/tunib-ai/tunib-electra](https://github.com/tunib-ai/tunib-electra)

## Stereotype-Detector

Korean stereotype detector with tunib-electra and k-stereoset (Nov 2021)

- [github.com/newfull5/Stereotype-Detector](https://github.com/newfull5/Stereotype-Detector)

## AI Poet-KoGPT2

Generative model for Korean poetry (Jun 2020)

- fine-tuning KoGPT2 from SKT-AI
- [github.com/newfull5/AI\\_Poet-KoGPT2](https://github.com/newfull5/AI_Poet-KoGPT2)

## PUBLICATION

---

### 하드웨어 초보자 가이드

오새찬 저 (사도출판)

eBook - basic knowledge about computer system (Oct 2019)

- [ridibooks.com/books/2773000027](https://ridibooks.com/books/2773000027)

## SKILLS & INTERESTS

---

- Programming Languages: Python, Shell Script, SQL
- Deep Learning: Pytorch, PyTorch-Lightning, Transformer
- MLOps: FastAPI, Flask, Docker

## AWARDS

---

- **Encourage Prize** - 2023 RoKAF Business Contest

## MILITARY SERVICE

---

### Republic of Korea Air force

Gimhae, Korea

Sergeant (May 2022 - Feb 2024)

- Network Control personnel in 5th Air Mobility Wing

