Sangchun Ha, Undergraduate Student

19, Misagangbyeondong-ro, Hanam-si, Gyeonggi-do, South Korea, +82-10-4725-9896, seomk9896@gmail.com

Date of birth 1997.11.07 Place of birth Gangdong-gu, Seoul

LINKS

GitHub, Linked In

PROFILE

Currently working as an intern at TUNiB, I developed the pre-trained language model, TUNiB Electra. And I am also responsible for detecting personal information, hate speech and sensitive topics.

I believe speech prompt technology and large-scale language model will be very important in the future. So, I am specially interested in both speech and natural language processing. If you are interested in me, please don't hesitate to contact me.

EMPLOYMENT HISTORY

Aug 2021 — Present

TUNiB, Internship

Gangnam, Seoul

Development Korean-English Bilingual Electra Models. [link]

Experienced in data collection and data preprocessing.

Researched services that detect personal information, hate speech, and sensitive topics.

Check more details about TUNiB at [link].

Mar 2021 — Aug 2021

NeuroAI Lab, Internship

Kwangwoon Univ.

Dept. of Biomedical Signal Processing & Human-Machine Interaction.

Studied end-to-end speech recognition technologies.

Check more details about NeuroAI Lab at [link].

EDUCATION

Mar 2016 — Present

Bachelor of Engineering, Kwangwoon University

Seoul, South Korea

Major in Electronic & Communication Engineering, Minor in Data Science.

Major GPA : 4.28 / 4.5, **Total GPA** : 4.12 / 4.5

Courses : AI & Speech Signal Processing, Data Mining, Text & Opinion Mining, Database, Capstone Design l-ll, Operating systems, Software Design, Computer Architecture, Data Structure & Algorithm, Object-Oriented Programming, Computer Network, Data Science, Big Data Planning and Analytics, C Programming, Data Communication, Digital Signal Processing, Signal and System, Digital Engineering, Digital Communication, Communication Theory, Network Analysis, Linear Algebra, Discrete Math, Engineering Mathematics l-ll, Physical Electronics, Circuit Theory, Semiconductor Devices, Electronic Circuit 1, Basic Electronic circuit Lab l-ll, Electronic Circuit LAB l-ll

PROJECT EXPERIENCE

Nov 2021 — Dec 2021

Safety

Developed services to identify and detect personal information, hate speech, and sensitive topics.

Experienced in data preprocessing with hate speech and personal information.

Aug 2021 — Sep 2021

TUNiB Electra

Released the pre-trained language model, TUNiB Electra. [link]

Experienced Electra model pre-training using TPU and fine tuning.

Experienced in data collection and data preprocessing.

May 2021 — Jul 2021

OpenSpeech

Implemented a framework to easily make a speech recognizer in various languages. [link]

Supports more than 20 speech recognition models including Transformer, Conformer, ContextNet, and Listen Attend Spell.

Dec 2020 — Feb 2021

Automatic Speech Recognition Models

Implemented End-to-End Speech Recognition models with PyTorch. [link]

Several speech recognition models: Listen-Attend-Spell, Deepspeech2, Speech-Transformer, RNN-Transducer, Transformer-Transducer, ContextNet

Developed three different attention mechanisms: Scaled dot-product attention, Location aware attention, Multi-head attention

AWARDS & HONORS

Oct 2021 — Nov 2021

AI Grand Challenge

3rd ranked, 2021 AI Grand Challenge - Speech Recognition Track

Apr 2021 — Jul 2021

CLOVA AI RUSH

Naver

4th ranked, Named Entity Recognition in Japanese order history. [link]
4th ranked, Extracting user embedding using large-scale shopping data.

Apr 2019 — Apr 2019

Software Start-Up Idea Contest

Kwangwoon University

Finalist, Kwangwoon University software start-up idea contest

EXTRA-CURRICULAR ACTIVITIES

Oct 2020 — Feb 2021

Google Machine Learning Bootcamp

Google Korea

Completed Deep Learning Specialization lecture which is taught by Prof. Andrew Ng.

Studying various deep learning models and techniques.

Obtaining the TensorFlow Developer Certificate. [link]

Oct 2021 — Nov 2021

Paper Reading

Composed of papers related to speech recognition and natural language

processing. [link]

SKILLS

Python	Expert	**C**	Experienced
PyTorch	Expert	**Git**	Skillful
PyTorch Lightning	Expert	**Matlab**	Skillful
huggingface	Expert	**SQL**	Skillful
Java	Experienced	**R**	Skillful

LANGUAGE PROFICIENCY

Aug 2020

OPIC IM

CERTIFICATE

Dec 2020 — Dec 2023

TensorFlow Developer Certificate, TensorFlow