

이력서



서덕진 남 1995년, 27세

휴대폰 | 010-3393-9581

Email | 406023@naver.com

전화번호 | 053-644-9502

주소 | 경기 용인시 수지구 풍덕천동

학력
금오공과대학교
대학교(4년)
졸업예정

경력
Clova AI Research, Naver Corp.
재직 중
총 10개월

인턴·대외활동 / 해외경험
논문 및 특허
외 3

자격증 / 어학
1종보통운전면허
OPIc IM2
외 1

학력

2014. 02 ~ 2021. 02
졸업예정
금오공과대학교 컴퓨터공학과
학점 | 3.11 / 4.5
졸업논문 | Deep Voice3 기반 음성합성 모델 개발

2014
졸업
대진고등학교

경력 총 10개월

2021. 04 ~ 재직중
Clova AI Research, Naver Corp. Clova Speech 인턴
대형 언어 모델 연구 개발 및 키워드 추출을 위한 Knowledge-Aware Language Model 연구
주요직무 | 딥러닝

2020. 09 ~ 2020. 12
4개월
디앤티랩스 데이터 바우처 사업
- 멀티 프로세서 기반 음성 데이터 구축 시스템 구축
- 음성인식 모델 및 음성인식 API 프로토타입 개발
주요직무 | 딥러닝, 데이터구축

2018. 06 ~ 2018. 10
5개월
케이앤랩 크롤러 개발
- 타겟 사이트 동적 크롤러 개발
주요직무 | 소프트웨어

인턴·대외활동

2020. 09 ~ 2020. 12
4개월
논문 및 특허 교내활동
Paper
- Transfer Learning from Speech Representations for Keyword Spotting (Submission scheduled)

- An Analysis of BERT Fine-tuning Approach for Keyphrase Extraction in Scientific Articles
- An Automatic Data Construction Approach for Korean Speech Command Recognition

Patent

- Voice data set building system based on subject domain

2018. 03 ~ 2020. 12
2년 10개월

교내 프로젝트 교내활동

- 음성 데이터 구축

웹 리소스로부터 대량의 음성 데이터를 구축하였습니다. 원하는 키워드에 대한 관련된 음성을 멀티 프로세서에 기반하여 획득하도록 설계하였습니다. 또한 음성인식 시스템을 구축하였습니다.

- 음성합성 모델 개발(학기 중 수업 프로젝트)

Deep Voice3과 Tacotron2를 통해 음성합성 모델을 제작하였습니다. 본 프로젝트에서는 한국어 음성 데이터를 유튜브에서 수집하여 음성합성 모델을 여러차례 제작해보려했지만 노이즈와 얼라이먼트의 문제로 제대로 합성이 되지 않았습니다. 마무리는 영어 음성합성 모델로 데모를 시연했지만, 한국어에서 아직도 해결해보고 싶은 문제입니다.

- 실시간 음성 명령어 인식

직접 음성 명령어 데이터셋을 구축하고, 앱 상에서 음성명령어 인식 모델을 개발하고 연동하였습니다. 모델의 성능을 높이기 위해 Clova TTS를 활용한 추가 데이터 구축 및 추가적인 전처리 기법을 사용하였습니다.

- 비트코인 시세 예측(창의설계프로젝트)

keras 프레임워크를 사용하여 Bi-LSTM 기반 비트코인 시세 예측 모델 개발

- KoQuAd(학기 중 수업 프로젝트)

Electra를 한국어 위키피디아+네이버 뉴스 데이터를 사용하여 사전 학습하고, KoQuAd task로 fine-tuning 한 경험이 있습니다. 활용할 수 있는 GPU환경이 제한적이어서 모델을 많이 학습해볼 수 없어 좋은 결과는 얻지 못 했지만, MLM의 구현에 대한 공부가 많이 되었습니다.

- 크롤러 개발

특정 사이트에서 타겟을 수집하여 DB화 하고 웹과 연동하는 작업을 한 경험이 있습니다.

2019. 09 ~ 2019. 10
2개월

공모전 사회활동

- Hodoo & Prediction STT Challenge

영어 발음기호를 예측하는 challenge에 참가하였습니다. 챌린지에 참가할 때, 논문 제출 기한이 얼마 남지 않아서 많은 시간을 할애하지는 못 했습니다. 저는 Wav2Vec 2.0을 기반으로 pretrained encoder를 사용하여 finetuning해서 validation에서는 3% 내의 WER을 획득하였으나, 본 챌린지에서 사전학습 모델은 사용하면 안되는 조건이 있어 아쉬운 참가였습니다.

- Clova Speech Hackathon

해커톤 참여 당시 음성인식을 구현하기에 실력이 많이 모자랐던 것 같습니다. 좋은 결과는 내지 못 했으나 여러 음성 처리 기법을 익힐 수 있는 좋은 기회였습니다.

2018. 09 ~ 2018. 12
4개월

멘토&조교 교내활동

- C언어 멘토링

교내 전공과목 멘토/멘티 프로그램

1학년을 대상으로 C언어 수업

- 안드로이드 프로그래밍 수업 조교
- 비 전공 대상 안드로이드 프로그래밍 수업 및 실습 조교

교육

2019. 10 ~ 2019. 10 **공공데이터 활용 분석/시각화 실습** 금오공과대학교

Python을 활용하여 공공데이터를 활용한 공공데이터 분석 및 시각화를 실습

자격증

2014. 07 **1종보통운전면허** 경찰청(운전면허시험관리단)

어학

영어

OPIc IM2 (취득일 : 2021. 01)

TOEIC 500점 (취득일 : 2020. 09)

자기소개서

지원직무에 대하여 자신이 알고 있고 응용할 수 있는 지식

Python

- 크롤러 개발 (Scrapy, Selenium, Splash)
- API 개발 (Django)

Pytorch

- Torch Lightning (DDP, DeepSpeed)
- End-to-End ASR (LAS, Deep Speech2, Jasper, QuartzNet, Transformer, Wav2Vec 2.0, Conformer)
- Meta Learning (Prototypical Net, Contrastive Learning)
- TTS (Deep Voice3, Tacotron2)
- Language Model (KenLM, ..., Bert, Electra)
- Key Phrase Extraction (based on NER, Keyphrase Generator)
- Keyword Spotting (MHA RNN, TC-ResNet)

Tensorflow

- Pratrained Model (Bart, Electra, Bert)
- Keyword Spotting (MHA RNN, TC-ResNet on Android)

Linux

- Docker

현재 연구 주제

- Knowledge-Aware Language Model
- Keyword Boosting

포트폴리오

포트폴리오	An Analysis of BERT Fine-tuning.pdf	포트폴리오	An Automatic Data Construction.pdf
포트폴리오	http://github.com/qute012		

희망근무조건

고용형태	정규직
희망근무지	서울전지역, 경기전지역
희망연봉	면접 후 결정
지원분야	직무 빅데이터AI(인공지능) > tensorflow, 머신러닝(ML) 직접입력 PyTorch

위의 모든 기재사항은 사실과 다름없음을 확인합니다.

작성자 : 서덕진

이 이력서는 2021년 04월 15일 (목)에 최종 업데이트된 이력서입니다.
위조된 문서를 등록하여 취업활동에 이용 시 법적 책임을 지게 될 수 있습니다.
잡코리아(유)는 구직자가 등록 한 문서에 대해 보증하거나 별도의 책임을 지지 않으며
첨부된 문서를 신뢰하여 발생한 법적 분쟁에 책임을 지지 않습니다.
또한 구인/구직 목적 외 다른 목적으로 이용시 이력서 삭제 혹은 비공개 조치가 될 수 있습니다.