

박영준

CONTACT INFORMATION	서울특별시 은평구 수색로 200	Voice: (+82)-10-5016-6133 E-mail: yeongjoon1227@gmail.com LinkedIn: https://www.linkedin.com/in/yeongjoonpark/
INTERESTS	자연어처리(NLP)와 Deep Learning 엔지니어입니다. 특히 자연어이해(NLU) , 대화 시스템 등에 관심이 많으며 새로운 기술을 적용하고 연구하는 것을 좋아합니다.	
CURRENT POSITION	NHN Diquest , 서울, 대한민국 전임연구원 AI 선형기술연구팀에서 전임연구원으로 근무중입니다. 자연어처리 관련 프로젝트들을 담당하고 있으며, Deep Learning 부분의 모델 개발 및 구현을 담당했습니다. 현재 전문연구요원 으로 2023년 4월까지 병역의무를 수행중에 있습니다.	Apr. 2020 - Present
EDUCATION	서강대학교, 컴퓨터공학과 M.S. 자연어처리 연구실 Advisor 서정연 교수 Research Area 자연어처리, 딥러닝	Mar. 2018 - Feb. 2020
	서강대학교, 컴퓨터공학과 B.S. 컴퓨터공학과 Cum Laude	Mar. 2014 - Feb. 2018
PROJECTS	패션 어드바이저 플랫폼 (2020 - Present) AI 기반의 패션 어드바이저 플랫폼에서 text tagging 부분을 담당. 패션 아이템의 설명 정보가 담겨있는 비정형 데이터에서 특성 정보들을 추출 하고 분류 후 API 형태로 제공. <ul style="list-style-type: none">• 데이터셋 구축, 데이터 전처리, 모델 학습, 데이터 후처리• BERT, ELECTRA, RoBERTa 등 사용	NHN Diquest
	ML 플랫폼 구축 (2022) Machine Learning 플랫폼 구축 프로젝트 중, 모델 설계/구현 부분을 담당. <ul style="list-style-type: none">• 데이터 전처리, 모델 학습,• Task: Dependency Parsing, Named Entity Recognition, Semantic Role Labeling	NHN Diquest
	민원 데이터 구축 (2021 - 2022) NLP 데이터 구축을 위한 프로젝트 중, 의도분석 부분을 담당. 딥러닝 모델을 이용한 데이터셋의 성능 지표 확보 및 정확도 개선. <ul style="list-style-type: none">• 데이터 전처리, 모델 설계/학습• BERT, ELECTRA 등 사용	NHN Diquest
	대화 시스템: Response selection (2019) DSTC8(Dialogue System Technology Challenge 8)에서 제공하는 데이터셋으로 End-to-End 대화 시스템 을 개발. 대화 내의 문장들간의 관계를 어텐션 매커니즘 을 활용해 표현하는 BERT 기반의 모델 개발.	서강대학교
	화행 분석 (2018 - 2019) 기계학습용 텍스트 데이터 레이블 자동생성 및 검증도구 개발 프로젝트의 화행 분석 부분을 담당. SVM 분류기 를 기반으로 bootstrapping 과 active learning 등의 학습 방법을 사용하여 텍스트 데이터의 화행 분석 모델 개발.	서강대학교
PUBLICATIONS	[ESWA '22] BERT-based Response Selection in Dialogue Systems Using Utterance Attention Mechanisms Yeongjoon Park , Youngjoong Ko, Jungyun Seo Expert Systems with Applications(ESWA), Volume 209, 2022	

[ICMLA '21] Medical Code Prediction from Discharge Summary: Document to Sequence BERT using Sequence Attention

Tak-sung Heo*, Yongmin Yoo*, **Yeongjoon Park***, Byeong-Cheol Jo*, Kyoungsun Kim

The 20th *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2021

***These authors contributed equally.**

[Symmetry '21] A Novel Hybrid Methodology of Measuring Sentence Similarity

Yongmin Yoo*, Tak-sung Heo*, **Yeongjoon Park***, Kyoungsun Kim

Symmetry, Aug 2021

***These authors contributed equally.**

[HCLT '19] 순차적 구문 분석 방법을 반영한 포인터 네트워크 기반의 한국어 의존 구문 분석기

한장훈, **박영준**, 정영훈, 이인권, 한정욱, 박서준, 김주애, 서정연

Proceedings of the 31st Annual Conference on Human and Cognitive Language Technology, pp. 533-536(4pages)

[KCC '19] 한국어 ELMo 임베딩을 이용한 의미역 결정

박찬민, **박영준**, 고영중, 서정연

Communications of the Korean Institute of Information Scientists and Engineers 46(01), 608-610(3pages)

[KSC '18] 화행 분석의 도메인 확장을 위한 부트스트래핑 기법이 적용된 계층적 SVM 분류기

박영준, 서민영, 김주애, 고영중, 서정연

Communications of the Korean Institute of Information Scientists and Engineers 45(02), 602-604(3pages)

[KCC '18] 전이 학습을 사용한 인공신경망 기반의 한국어-영어 기계번역

박영준, 안휘진, 김주애, 서정연

Communications of the Korean Institute of Information Scientists and Engineers 45(01), 595-597(3pages)

[KSC '17] 중복 질의 판단을 위한 기계 학습

김민정, **박영준**, 임형석, 양지훈

Communications of the Korean Institute of Information Scientists and Engineers 44(02), 1986-1988(3pages)

[KSC '17] 텍스트 매핑을 이용한 스케치 기반의 얼굴 이미지 생성

김민정, 임형석, **박영준**, 주예슬, 구명환

Communications of the Korean Institute of Information Scientists and Engineers 44(02), 2001-2003(3pages)

**HONORS AND
AWARDS**

ETRI Fashion How Season2 (2021) - 1st place

ETRI

시스템과 사용자의 대화 내용과 의상 이미지 및 메타 정보를 고려하여 사용자의 발화 내용에 적합한 의상코드를 추천.

- 복합모달(Multimodal)
- ELECTRA, ResNet 등 사용

Animal Datathon Korea (2021) - 2nd place

한국축산테크협회

소 이미지의 관절 좌표 예측 및 pose estimation.

ETRI Fashion How Season1 (2020) - 1st place

ETRI

시스템과 사용자의 대화 내용과 의상 메타 정보를 고려하여 사용자의 발화 내용에 적합한 의상코드를 추천.

- ELECTRA 사용

국어 정보 처리 시스템 경진대회 (2019) - 1st place

국립국어원

한국어 의존구문 분석

Naver Challenge (2018)

NAVER

한국어 의미역결정(SRL) - **1st place**

한국어 개체명인식(NER) - **4th place**

**TEACHING
EXPERIENCE**

AI 전문가 과정: 자연어처리

삼성전자

실습 조교

Jun - Nov. 2019

삼성전자 AI 전문가 과정에서 실습 조교로 근무. **Word2vec, RNN, Attention, Transformer, BERT** 등 자연어처리 전반적인 내용 및 대화 시스템에 대해 중점적으로 강의 및 실습 진행. DSTC(Dialog System Technology Challenge) 7 데이터셋으로 response selection 실습 진행.

SKILLS

- **Languages:** Python, C, C++,
- **Frameworks:** Pytorch, TensorFlow/Keras
- **Tools:** Docker, GIT
- **Others:** VSCode, Vim, Pycharm