

## EDUCATION

<b>Carnegie Mellon University</b> MASTER IN LANGUAGE TECHNOLOGIES — <i>Advisor: Eduard Hovy</i>	Sep. 2022 - Current Pittsburgh, USA
<b>Carnegie Mellon University</b> VISITING STUDENT — <i>Funded by Korean Government</i>	Sep. 2019 - Feb. 2020 Pittsburgh, USA
<b>Sogang University</b> M.S. IN COMPUTER SCIENCE AND ENGINEERING — <i>CGPA: 4.38 / 4.5</i>	Mar. 2018 - Feb. 2021 Seoul, Korea
<b>Sogang University</b> B.A. IN AMERICAN CULTURE, PSYCHOLOGY — <i>Summa Cum Laude</i> 2-year leave of absence as military interpreter	Mar. 2011 - Feb. 2018 Seoul, Korea

## EXPERIENCE

<b>NLP Research Engineer</b> NAVER, <a href="#">CLOVA</a> <ul style="list-style-type: none"><li>Implemented efficient transfer learning methods using model and data parallelism, enabling finetuning GPT3 to downstream tasks</li><li>Researched effect of pretraining corpora on in-context learning of GPT-3 and coauthored NAACL 2021 paper</li><li>Developed python library for pretraining and distributing language models which is used company-wise</li></ul>	Jan. 2021 - May. 2022 Seongnam, Korea
---	--

## PROJECTS

<b>Measuring the impact of AI on U.S economy</b> CARNEGIE MELLON UNIVERSITY, LANGUAGE TECHNOLOGIES INSTITUTE <ul style="list-style-type: none"><li>Developing BERT based classifier tailored for patent data classification, discovering AI related patents</li><li>Extracting AI experts from publication dataset with LDA to track their trajectory</li></ul>	Aug. 2022 - Current Pittsburgh, USA
<b>Grammar error correction</b> SOGANG UNIVERSITY, NATURAL LANGUAGE PROCESSING LAB <ul style="list-style-type: none"><li>Proposed method to generate diverse synthetic data using self training and back translation to increase BLEU score by 5.26</li></ul>	Jul. 2020 - Sep. 2020 Seoul, Korea
<b>Quantifying cross-cultural similarity</b> CARNEGIE MELLON UNIVERSITY, LANGUAGE TECHNOLOGIES INSTITUTE <ul style="list-style-type: none"><li>Proposed cultural similarity features for cross-lingual transfer that can be used to rank transferable languages</li></ul>	Dec. 2019 - Jul. 2020 Pittsburgh, USA
<b>Multilingual offensive language detection</b> CARNEGIE MELLON UNIVERSITY, LANGUAGE TECHNOLOGIES INSTITUTE <ul style="list-style-type: none"><li>Devised cross-lingual data selection method for offensive language detection and achieved competitive results in SemEval 2021 including 1st place in Arabic</li></ul>	Dec. 2019 - Feb. 2020 Pittsburgh, USA
<b>Dependency Parsing</b> SOGANG UNIVERSITY, NATURAL LANGUAGE PROCESSING LAB <ul style="list-style-type: none"><li>Incorporated graph based and transition based parser and examined the role of morphemes, accomplishing SOTA in Korean dependency parsing and applied for patent</li></ul>	Jul. 2019 - Sep. 2019 Seoul, Korea

## SKILLS

<b>Programming Languages:</b> Python, C
<b>ML libraries:</b> PyTorch, Transformers, Datasets, Pytorch Lightning, MegatronLM, Numpy, Pandas, Scikit-learn
<b>Dev environment:</b> Git, Docker, Vim, Tmux, Shell, Linux

## PUBLICATIONS

Seongjin Shin*, Sang-Woo Lee*, Hwiyeen Ahn, Sungdong Kim, HyoungSeok Kim, Boseop Kim, Kyunghyun Cho, Gichang Lee, Woomyoung Park, Jung-Woo Ha, Nako Sung. On the Effect of Pretraining Corpora on In-context Few-shot Learning by a Large-scale Language Model. <i>NAACL 2022</i> . <a href="#">[pdf]</a>
Hwiyeen Ahn*, Jimin Sun*, Chan Young Park*, Yulia Tsvetkov, David R. Mortensen. Cross-Cultural Similarity Features for Cross-Lingual Transfer Learning of Pragmatically Motivated Tasks. <i>EACL 2021</i> <a href="#">[pdf]</a>
Hwiyeen Ahn*, Jimin Sun*, Chan Young Park*, Jungyun Seo. NLPDove at SemEval-2020 Task 12: Improving Offensive Language Detection with Cross-lingual Transfer. <i>SemEval 2020</i> . <a href="#">[pdf]</a>
Hwiyeen Ahn, Minyoung Seo, Chanmin Park, Juae Kim, Jungyun Seo. Extensive Use of Morpheme Features in Korean Dependency Parsing. In <i>IEEE BigComp 2019</i> .