

Language and Data Intelligence Lab

Seung-won Hwang

Professor, SNU CSE

March 12, 2021

Principal Investigator: Seung-won Hwang

- ▶ As a professor (since 2005), areas of interests have been Natural Language Processing, Information Retrieval, and Data Mining.
- ▶ 150+ papers published, targeted to conferences: ACL, EMNLP, NAACL, EACL, SIGIR, WSDM, AAI, VLDB, SIGMOD, ICDE; Graduated 12 PhDs and 16 MS
- ▶ KAIST (BS) UIUC (MS/PhD)
- ▶ Further details can be found at:
 - ▶ seungwonh@snu.ac.kr
 - ▶ <http://seungwonh.github.io>
 - ▶ 301-520 (by email appointment)

Showcase 1: Knowledge injection for neural NLP models

- ▶ Our recent NAACL paper injects symbolic knowledge (from Wikipedia KB) to augment an intermediate neural representation in the neural models.
- ▶ Injecting other modalities of knowledge, such as translational knowledge for our IJCAI paper.

Showcase 2: Self-training neural NLP models

- ▶ Neural models are accurate, but being optimized for accuracy alone, may predict “right answer for wrong reason” (not robust!).
- ▶ We mine human supervisions (without asking humans to annotate) to improve the robustness of neural models (SIGIR, EMNLP, and ICDE papers)

Ongoing: Sustainable Training of Language Models

- ▶ Language models are being trained with hundreds of billions of parameters, and not only this massive-scale pre-training, but also fine-tuning pre-trained models with task-specific annotation, require lots of computational and annotational overhead.
- ▶ As a complementary direction of building more and more complex models, we are replacing computational/annotational overhead with the mined knowledge, which is especially important for lower-resource languages.

Students and Alumni

- ▶ Current: 8 PhD, 4 MS Students
- ▶ Graduated: 12 PhD. 16 MS Students
- ▶ Alumni are employed as assistant professors (Yonsei and SKKU), researchers (IBM Research, Facebook, Argonne National Lab, Samsung Research, Naver, NCSOFT AI, POSCO), and a start-up founder (HyperConnect).