Human Knowledge for Robustifying AI

Seung-won Hwang*

Department of Computer Science and Engineering,
Seoul National University
Seoul, Korea
**Presenting/Corresponding Author: seungwonh@snu.ac.kr*

Deep learning models for NLP often tackle each task in isolation, requiring a large number of training examples and works well only on well-defined and narrow tasks. Meanwhile, we face the challenge of robustness, e.g., when generalizing to poor-resource languages or out-of-distribution applications. This abstract is for my 20-min talk, overviewing our recent approaches of using human knowledge for robustifying AI models, specifically using counterfactual knowledge (EMNLP 2020) and augmentation (ACL 2021). Further details can be found at: https://seungwonh.github.io/