

A local approach for improving isotropy in contextual embedding space

سارا رجایی، پژوهشگاه دانش‌های بنیادی

Abstract

The representation degeneration problem in Contextual Word Representations (CWRs) hurts the expressiveness of the embedding space by forming an anisotropic cone where even unrelated words have excessively positive correlations. Existing techniques for tackling this issue require a learning process to retrain models with additional objectives and mostly employ a global assessment to study isotropy. Our quantitative analysis over isotropy shows that a local assessment could be more accurate due to the clustered structure of CWRs. Based on this observation, we propose a local cluster-based method to address the degeneration issue in contextual embedding spaces. We show that in clusters including punctuations and stop words, local dominant directions encode structural information, removing which can improve CWRs performance on semantic tasks. Moreover, we find that tense information in verb representations dominates sense semantics. We show that removing dominant directions of verb representations can transform the space to better suit semantic applications. Our experiments demonstrate that the proposed cluster-based method can mitigate the degeneration problem on multiple tasks.

Biography

Sara Rajaei is an M.Sc. student in artificial intelligence at Iran University of Science and Technology (IUST), under the supervision of Prof. Mohammad Taher Pilehvar. Her research interest mainly revolves around the analysis and interpretability of NLP models. More specifically, she is interested in deep language models' ability to understand and comprehend human languages.

زمان: یکشنبه ۲۰ تیر ۱۴۰۰، ساعت ۱۸

ارائه بصورت مجازی انجام خواهد شد.

<https://cp.fararoom.ir/ch/ipm/guest>

*** شرکت برای عموم علاقه‌مندان آزاد است ***