

News & Comments

Chinese Short Text Entity Linking Method Based on Siamese-Like Network

Jenson Easo

Entity Linking (EL), which is frequently used as pre-processing for tasks like information extraction and intelligent question and answer, is regarded as an essential component of natural language processing. By connecting various texts with structured data, it also plays a crucial part in the creation and updating of knowledge bases. Since English does not have word segmentation restrictions and that there are numerous entities linking knowledge bases, entity linking technology has currently progressed greatly in this language. Entity linking can be viewed as a semantic matching issue between the text describing the candidate entity and the text mentioning it. Machines can now comprehend text content more and more effectively thanks to the development of pre-training models. To the disadvantage of not fully utilizing the text information of the entity to be disambiguated, the similarity between the entity description text to be disambiguated and the candidate entity description text is determined directly through a pre-training model for the disambiguation aspect. A key method for addressing the issue of words with multiple meanings and multiple words with the same meaning is entity linking. Entity linking involves processing both the text's contents and the knowledge base's contents to match the entities in the text that need to be disambiguated with the knowledge base.

The method for linking entities that is suggested in this research is based on a Siamese-like network and differs from the conventional approach, which simply measures the similarity between entity mentions and entities. The semantic relationship in the text is thoroughly analysed in this study, not only by feature representation of the disambiguated entity text features and candidate entity text features, but also by making full use of the text features of the entity to be disambiguated to extract the most important information. The experimental results demonstrate that, as compared to the baseline model, the method presented in this study achieves higher accuracy in Chinese short textual entity linking. Future research will look at how to translate the current approach for multimodal entity linking and entity linking difficulties to other languages. In the actual world, data is more often presented in an "image + sound + video" format, combining the visual and auditory modalities. Due to lacking data, entity linkage that just takes into consideration textual information may be erroneous.

Source: [Information](#)

KEYWORDS

Chinese short text, entity linking, named entity recognition, entity disambiguation

