

Novel Method to Encode Sentences for Downstream Tasks

Description:

Sentence encoder is the de facto way to transform a sequence of words into a latent representation. The most common architecture is the *sequence-to-sequence* [1] which uses a LSTM as encoder, an attention mechanism in the latent space and another LSTM as decoder. Other works propose the use of (normal or dilated) convolutions [2, 3] relying on consecutive n-grams. Finally, another kind of architecture called *Transformer*, based purely on self-attentions, has been proposed in [4]. All of these have different pros and cons but none exploits additional information among words besides word embeddings.

In this project, we would like to experiment other ways to encode sentences by using graph representations. Moreover, there exist various ways to model sentences as graphs. As models, we could apply on top graph convolutional neural networks [5] or even graph attention networks [6].

Once different sentence encoders have been developed, we could evaluate their performance with different models available in the literature on different downstream NLP tasks such as sentiment analysis, textual entailment, summarization and many more.

Prerequisites: Knowledge about Machine Learning & efficient with Python

Supervisor: Diego Antognini (diego.antognini@epfl.ch)

References:

- [1] <https://arxiv.org/pdf/1409.0473.pdf>
- [2] <https://arxiv.org/pdf/1612.08083.pdf>
- [3] <https://arxiv.org/pdf/1705.03122.pdf>
- [4] <https://arxiv.org/pdf/1706.03762.pdf>
- [5] <https://arxiv.org/pdf/1609.02907.pdf>
- [6] <https://arxiv.org/pdf/1710.10903.pdf>