Is It Worth the Attention?

A Comparative Evaluation of Attention Layers for Argument Unit Segmentation

Maximilian Spliethöver ■
Jonas Klaff ■
Hendrik Heuer ■

The general approach tackles the unit segmentation problem for argumentative texts. It is based on an approach by Ajjour et al. [1], with the aim to enhance the performance, as well as simplifying the preprocessing by using different input features in addition to the added attention layer. The concept of adding attention to the architecture is derived from recent advances in Neural Machine Translation [2,3]. The performance is evaluated on the "Argument annotated essays" corpus compiled by Subh, et al. [6].

An attention-based pre-trained language model [5], derived from the Encoder part of the Transformer [2] architecture. Embeddings are generated by feeding one sentence into the language model and extracting the representations for the first subword token from the last four layers of the pre-trained BERT model via the Flair library. The subtoken embedding is then used as representations for the whole token.

A pre-trained character-level Bi-LSTM [4] to predict the most probable next character for a given input sequence. Embeddings are generated by feeding one sentence into the language model, extracting the corresponding weights via the Flair library and stack them with GloVe300 embeddings.

* calculates a relevance score for each part of an input sequence
* self-attention: with respect to the complete input vector of a single token
* multi-head attention: with respect to distinct sub-spaces of the token's input vector

<table>
<thead>
<tr>
<th>architectures</th>
<th>baseline</th>
<th>baseline with attention</th>
<th>single bi-LSTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>embedding</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>GloVe</td>
<td>0.86</td>
<td>0.85 0.67</td>
<td>0.86 0.84</td>
</tr>
<tr>
<td>BERT</td>
<td>0.83</td>
<td>0.68 0.68</td>
<td>0.86 0.83</td>
</tr>
<tr>
<td>Flair</td>
<td>0.87</td>
<td>0.67 0.67</td>
<td>0.86 0.81</td>
</tr>
</tbody>
</table>

REFERENCES