

What does Attention in Neural Machine Translators do?

Friday, 15 November 2019

7:17 PM

The paper studies cases where word alignment does not follow attention. Authors check to see if this is a desirable behaviour of the attention model or error.

- Have alignment as the ground truth from the dataset - Get soft alignment
- Get attention loss - is low when - alignment is high only when attention is high
 - If alignment has a high score for low attention, then the loss value increases
- Also, just getting a low attention loss is not enough for saying that attention is good. We also need to see how the NMT system performs for the values of alignment.
- Hence, We measure the Spearman's rank correlation between the system performance and alignment loss.
- Authors **chose not to use BPE (Sennrich et al., 2016) which operates on tokens**
- Authors report that attention to other parts of the sentence (and not just the current word) is essential for translating some words (Predicting POS tag of verbs in the experiment)
- While for translating Nouns - it is attention and alignment must be similar
 - These results are inferred using the Spearman Correlation between attention loss and prediction loss.
 - For verbs, this correlation is low indicating high attention loss corresponds to high prediction loss, but for nouns this correlation is high - signaling high attention loss corresponds to low prediction loss
- Authors also measure the concentration of attention using : Entropy of attention weights
- the attention is concentrated in case of nouns while it is dispersed in case of verbs

Spearman's correlation constant basically measures how well the relationship between two variables can be described using a monotonic function.

translation pay

error on it's part.

h.

ases.

and alignment are the

attention loss.

formance and the attention

at the sub-word level

correct alignment) is

ment).

ntion loss and the word

ding with low prediction

matching with high word

ion.

verbs.

tween two

~

GIZA++ --> A tool for bilingual sentence alignment

<https://medium.com/@weGotlieb/using-giza-for-bilingual-sentence-alignment-8b1e1e1e1e1e>

[b2dc828a04d](#)