Skeleton-based Coherence Modeling in Narratives

Nishit Asnani, Rohan Badlani

Department of Computer Science, Stanford University

Abstract

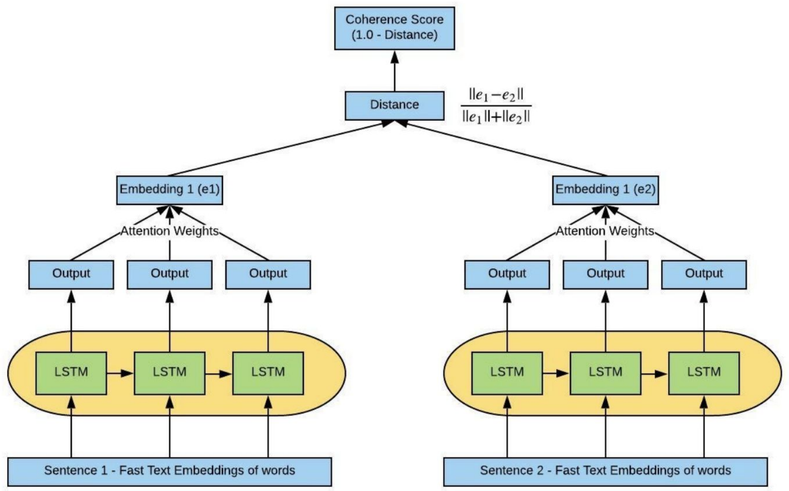
Analyze whether skeletons (or key phrases) from sentences within a given text sequence can be used for coherence modeling. We propose a new architecture called SSN that can capture similarity between sentences or skeletons. Although SSN is better than non-parametric similarity measures but sentences are much better than skeletons for coherence.

Introduction

Modelling coherence is a challenging task since even humans are not perfect at it. There have been quite a few approaches to modelling coherence and some of them include entity-grid representations, domain independent models & intentional structure in coherence assessment. Motivated by the progress in this direction, through this project we test if skeletons or key-phrases within a text is a good way to measure the coherence of the text.

Proposed Network

Sentence/Skeleton Similarity Network (SSN)



SSN Siamese Network Contrastive Loss

$$E_w = \frac{cosine(e_1, e_2)}{|\ e_1\ ||\ e_2\ |} \quad L_{pos}(e_1, e_2) = \frac{1}{4} * (1 - E_w)^2 \quad L_{neg}(e_1, e_2) = \begin{cases} E_w^2, \text{if} E_w < m \\ 0, else \end{cases}$$

$$L_w^i(e_1^i, e_2^i, y^i) = y^i L_{pos}(e_1^i, e_2^i) + (1 - y^i) L_{neg}(e_1^i, e_2^i)$$

Attention

$$c_i = \sum W_{ij} o_j$$

$$\alpha = softmax(c)$$

$$o' = \sum \alpha_i o_i$$

Data

- Skeleton Extraction: We use the story telling dataset to get both sentences and skeletons that are coherent and evaluate the performance of proposed SSN model
  - 40153, 4990, and 5054 stories for training, validation, and testing.
  - Maximum 6 sentence long
- Skeleton/Sentence Similarity Network: We construct 2 sets of datasets for SSN network:
  - Consecutive Sentence Similarity Evaluation: we pick 2 consecutive sentences within story as similar and across stories as dissimilar
  - Story Similarity Evaluation: Prepare the ordered and jumbled set

Methods

- **Skeleton Extraction**: We use input-to-skeleton model proposed by Jingjing Xu et. al at EMNLP 2018 to generate skeletons for a given sentence.
- **FastText Embeddings**: We train a Facebook research FastText embeddings model for our data to get word vector embeddings
- **Sentence/Skeleton Similarity Network (SSN)**: We train the above proposed model for each of sentences and skeletons and evaluate the performance based on 3 metrics
  - Sentence-order accuracy: Given a pair of consecutive sentences and a pair of randomly sampled sentences, is the model able to distinguish between the two?
  - Story-order accuracy: Given the original story as a list of sentences and a copy of the same story with sentence order randomized, is the model able to distinguish the two?
  - Pair-classification accuracy: Given two sentences, is the model able to detect if they are consecutive (by implicitly measuring coherence) or not?

Results

Comparing SSN with Baselines

Model-Type	Technique	Sentence-order Accuracy
Sentence	BERT Emb + Euclidean Distance	68.30%
Sentence	BERT Emb + Cosine Similarity	71.90%
Sentence	FastText Emb + 3-layer SSN without Attention	92.90%
Sentence	FastText Emb + 2-layer SSN with Attention	92.30%
Skeleton	BERT Emb + Euclidean Distance	59.30%
Skeleton	BERT Emb + Cosine Similarity	61.60%
Skeleton	FastText Emb + 3-layer SSN without Attention	84.20%
Skeleton	FastText Emb + 2-layer SSN with Attention	84.50%

Comparing Sentences with Skeletons

Model-Type	Technique	Sentence-order	Story-order	Sentence-Pair
Sentence	FastText Emb + 3-layer SSN without Attention	92.90%	69.60%	82.20%
Sentence	FastText Emb + 2-layer SSN with Attention	92.30%	68.00%	81.40%
Skeleton	FastText Emb + 3-layer SSN without Attention	84.20%	62.90%	73.80%
Skeleton	FastText Emb + 2-layer SSN with Attention	84.50%	62.30%	74.50%

Conclusions

- Neural approaches like SSN to modeling coherence work better than non-parametric approaches even after using BERT embeddings.
- Empirical results show that the sentence based techniques perform better than skeletons in the task of coherence detection between 2 sentences.
- Both sentence-based and skeleton-based SSN models perform much better at sentence-level coherence detection than paragraph-level coherence detection.
- 2-layered LSTM model with attention gives almost similar performance as the 3-layered LSTM model without attention on all metrics. It is expected that with 3-layered LSTM with attention will perform even better.

Future Scope & Extensions

- Evaluation on longer length datasets will provide a more conclusive result in terms of coherence detection on larger sequences of text
- We applied a very simplistic self-attention mechanism in this paper which can be extended to complicated ideas like the ones presented in Transformers.
- Evaluating the proposed SSN in detecting incoherent sentences within a text document with 1-2 sentences incoherent rather than fully ordered or jumbled stories, essays or paragraphs.