AQUAMUSE: Release notes V2

About this release

- Includes all query-long answer pairs from Natural Questions dataset that satisfy the criteria discussed in the paper `https://arxiv.org/pdf/2010.12694.pdf`.

Data set statistics

		# examples			summary		inputs		per-input doc	
Dataset	# queries	train	dev	test	# words	# sents	# words	# sents	# words	# sents
AQUAMUSE	7,725	6,253	661	811	107.8	3.7	9744.2	390.6	1601.5	64.2

Table 1: Number of examples and average inputs and summary sizes in this release.

Baselines

Method	R-1	R-2	R-L
Query-agnostic setting			
PEGASUS Zhang et al. (2019a)	25.2	6.78	12.86
Query-based setting			
PEGASUS	24.62	5.78	17.16

Table 2: Abstractive baseline evaluation on test split

Method	R-1	R-2	R-L
Query-agnostic setting			
NeuSum Zhou et al. (2018)	45.83	34.69	44.38
HIBERT Zhang et al. (2019b)	44.2	28.71	32.08
TextRank Mihalcea and Tarau (2004)	24.4	15.56	31.6
Query-based setting			
NeuSum	47.40	36.13	46.27
HIBERT	44.96	29.73	32.96
TextRank	25.72	17.4	34.3

Table 3: Extractive baseline evaluation on test split

References

Rada Mihalcea and Paul Tarau. 2004. https://www.aclweb.org/anthology/W04-3252 TextRank: Bringing order into text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pages 404–411, Barcelona, Spain. Association for Computational Linguistics.

Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter J. Liu. 2019a. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. *ArXiv*, abs/1912.08777.

Xingxing Zhang, Furu Wei, and Ming Zhou. 2019b. http://arxiv.org/abs/1905.06566 HIBERT: document level pre-training of hierarchical bidirectional transformers for document summarization. *CoRR*, abs/1905.06566.

Qingyu Zhou, Nan Yang, Furu Wei, Shaohan Huang, Ming Zhou, and Tiejun Zhao. 2018. https://doi.org/10.18653/v1/P18-1061 Neural document summarization by jointly learning to score and select sentences. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 654–663, Melbourne, Australia. Association for Computational Linguistics.