Appendix

Frédéric Piedboeuf
RALI, Diro
Université de Montréal
frederic.piedboeuf@umontreal.ca

Philippe Langlais
RALI, Diro
Université de Montréal
felipe@iro.umontreal.ca

	Best	**T5-f**	**T5-e**	**T5-m**	**T5-a**
Papyrus-f	**30.4/3.6**	20.1/0.6	12.1/0.2	20.0/0.4	18.8/0.6
Papyrus-e	**34.6/4.3**	2.8/0	24.3/0.5	24.1/0.7	15.6/0.3
Papyrus-m	**30.6/3.4**	11.9/0.3	18.2/0.3	22.0/0.6	17.1/0.5
Papyrus-a	**22.6/3.0**	13.4/0.4	13.4/0.4	12.4/0.4	13.4/0.3
Wikinews (fr)	**27.4**/0.5	20.4/0.3	7.0/0	21.8/**1.2**	16.4/0.3
cacis57 (es)	**48.0/0**	21.6/0	19.1/0	27.4/0	13.6/0
wicc 78 (es)	**30.1/0**	13.1/0	12.4/0	20.4/0	11.2/0
110ptbnkp (pt)	**21.4/0**	6.6/0	4.2/0	8.9/0	5.4/0
pak2018 (pl)	**7.4/4.6**	0/0	0.3/0	0.5/0	0/0
Inspec	**28.5/1.4**	2.3/0	19.3/0.3	20.7/0.3	9.8/0.1
NUS	**35.7/3.0**	3.5/0	24.8/1.0	24.4/1.4	15.9/0.5
SemEval	**22.9/1.6**	2.2/0	17.1/0.2	15.6/0.4	10.6/0.1
Kp20K	**30.5/5.0**	1.1/0	17.6/0.4	17.1/0.5	10.5/0.2
Krapivin	**28.0/5.8**	1.3/0	15.9/0.5	15.7/0.6	10.0/0.2

Table 1: F1@10 present/R@10 absent, on all sets for T5. Best represents the best results reported by any system in the paper. We notice that T5 underperforms, getting better R@10 only on the Wikinews dataset. We found that T5 was faster to converge but overfitted easily.

Submitted to the 36th Conference on Neural Information Processing Systems (NeurIPS 2022) Track on Datasets and Benchmarks. Do not distribute.