



Contrastive Aligned Joint Learning for Multilingual Summarization

Danqing Wang, Jiaze Chen, Hao Zhou, Xipeng Qiu †,
Lei Li

ByteDance AI Lab

†Fudan University



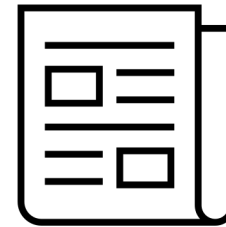
ByteDance AI Lab
字节跳动人工智能实验室



A New Multilingual Summarization Dataset: MLGSum

Total 1.1million articles in 12 langauges

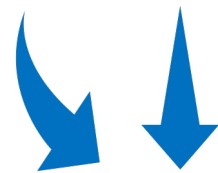
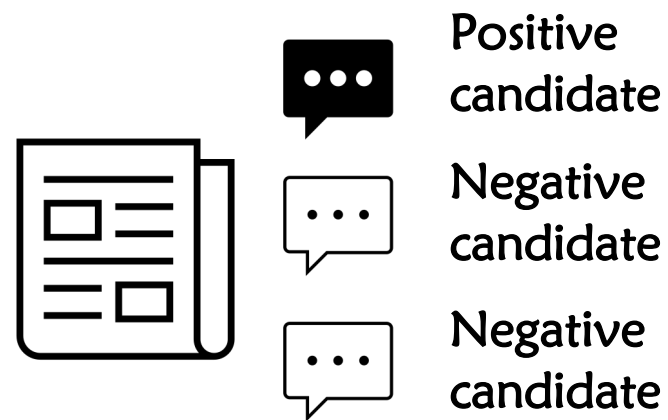
Language	Size	Article	Summary	Train
De	494,514	457	27	445,062
En	191,365	476	24	172,228
Ru	87,125	499	24	78,412
Fr	85,030	463	36	76,527
Zh	65,203	799	56	58,682
Hi	59,145	565	28	53,230
Es	43,162	703	30	38,845
Id	35,495	360	21	31,945
Tr	26,539	342	20	33,047
Vi	26,539	847	34	23,885
Uk	33,214	444	21	29,892
Pt	20,945	927	34	18,850
Total	1,168,276	573.5	29.6	1,060,605



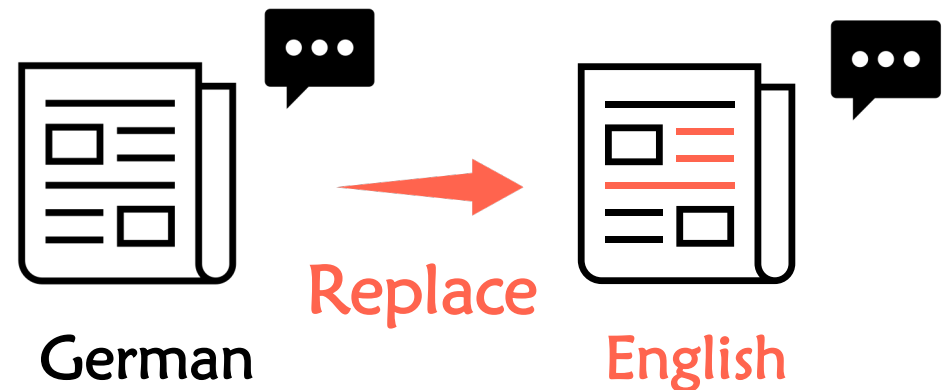
[Fr] Arsenal, le leader de la Premier League, a é té sé ve` rement corrigé, samedi 14 dé` cembre, par Manchester City (6-3) qui prend la deuxie` me place du classement a` seule- ment trois points des hommes d'Arse` ne Wenger. [q] Il s'agit de la huitie` me victoire de City a` domicile ou` il est invaincu cette saison, et ce contre la meilleure é` quipe a` l'exté` rieur.

[Fr] Irré` sistent a` domicile depuis le de` but de la sai- son, Manchester City a é` trillé` Arsenal (6-3) lors du match au sommet de la Premier League. [q] Les Mancuniens revi- ennent a` trois points des Gunners en haut du classement.

CALMS: Contrastive Aligned Learning

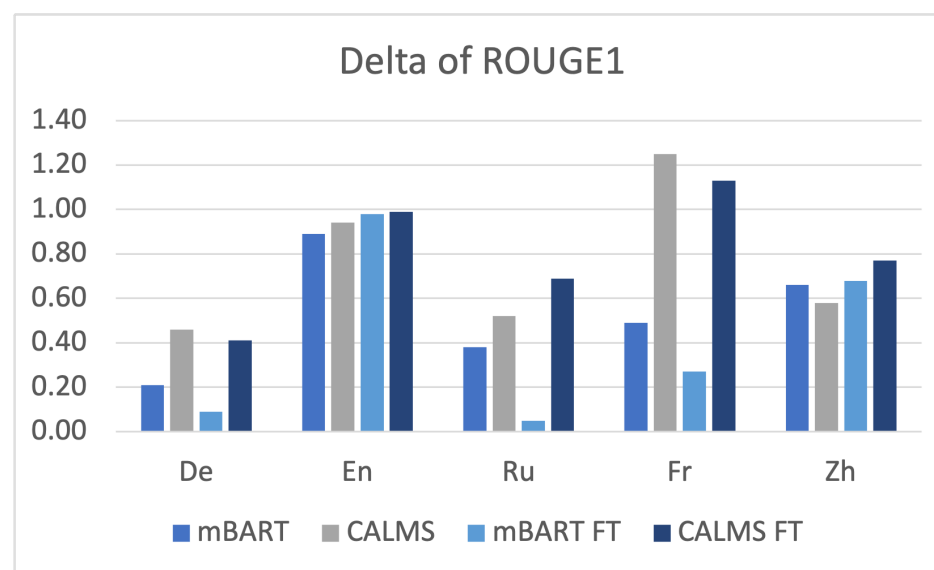
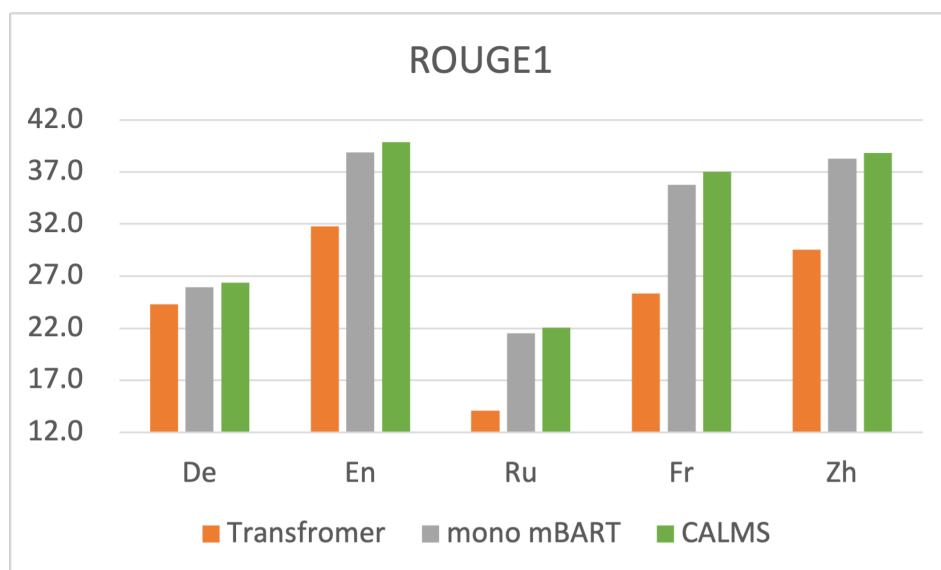


Contrastive Sentence Ranking



Sentence Aligned Substitution

Outperform monolingual models for each language



Thanks for your listening !

Contrastive Aligned Joint Learning for Multilingual Summarization

