

## Benchmark results

*The results of our benchmark for several Language Models using data from MMS.*

Our preliminary results has been presented in (Rajda et al. 2022) and finally presented in (Augustyniak et al. 2023) review at NeurIPS'23.

### Benchmark results - F1 Macro scores

#### Models

Model	Inf. time [s]	#params	#langs	base	data	reference
mT5	1.69	277M	101	T5	<i>CC</i> <sup>b</sup>	(Xue et al. 2021)
LASER	1.64	52M	93	BiLSTM	<i>OPUS</i> <sup>c</sup>	(Artetxe and Schwenk 2019)
mBERT	1.49	177M	104	BERT	Wiki	(Devlin et al. 2019)
MPNet**	1.38	278M	53	XLNet	<i>OPUS</i> <sup>c</sup> , <i>MUSE</i> <sup>d</sup> , <i>Wikititles</i> <sup>e</sup>	(Reimers and Gurevych 2020)
XLNet-dist**	1.37	278M	53	XLNet	<i>OPUS</i> <sup>c</sup> , <i>MUSE</i> <sup>d</sup> , <i>Wikititles</i> <sup>e</sup>	(Reimers and Gurevych 2020)
XLNet	1.37	278M	100	XLNet	CC	(Conneau et al. 2020)
LaBSE	1.36	470M	109	BERT	CC, Wiki + mined bitexts	(Feng et al. 2020)
DistilBERT	0.79	134M	104	BERT	Wiki	(Sanh et al. 2020)
mUSE-dist**	0.79	134M	53	DistilBERT	<i>OPUS</i> <sup>c</sup> , <i>MUSE</i> <sup>d</sup> , <i>Wikititles</i> <sup>e</sup>	(Reimers and Gurevych 2020)
mUSE-transformer*	0.65	85M	16	transformer	mined QA + bitexts, SNLI	(Yang et al. 2020)

Model	Inf. time [s]	#params	#langs	base	data	reference
mUSE-cnn*	0.12	68M	16	CNN	mined QA + bitexts, SNLI	(Yang et al. 2020)

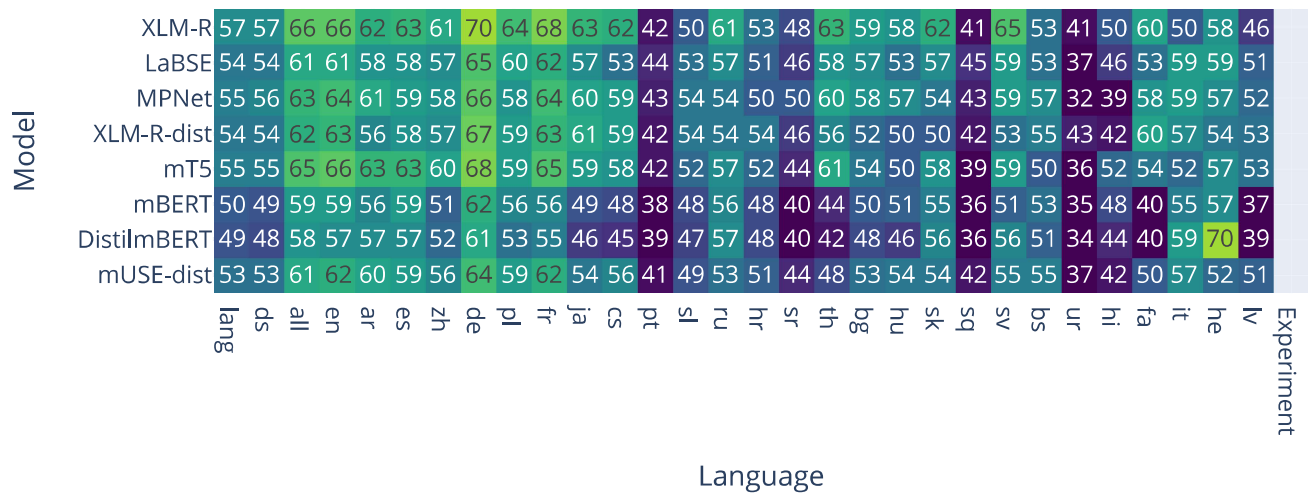
- \* mUSE models were used in TensorFlow implementation in contrast to others in torch
- a Base model is either monolingual version on which it was based or another multilingual model which was used and adopted
- b Colossal Clean Crawled Corpus in multilingual version (mC4)
- c multiple datasets from OPUS website (<https://opus.nlpl.eu>)
- d bilingual dictionaries from MUSE (<https://github.com/facebookresearch/MUSE>)
- e just titles from wiki articles in multiple languages

## Results

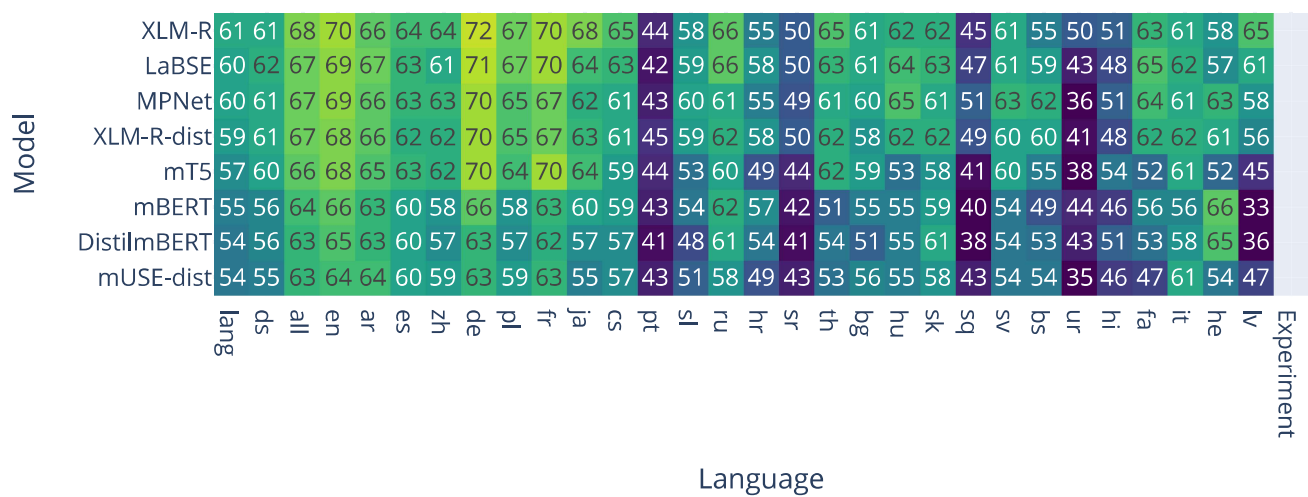
### Linear Head

Model	Experiment																															
	lang	ds	all	en	ar	es	zh	de	pl	fr	ja	cs	pt	sl	ru	hr	sr	th	bg	hu	sk	sq	sv	bs	ur	hi	fa	it	he	lv		
XLM-R	52	51	61	58	48	59	59	67	58	63	61	58	39	49	53	45	48	58	49	53	53	38	56	53	41	45	45	55	54	41		
LaBSE	54	53	61	59	53	60	57	65	58	63	60	59	44	55	55	52	48	55	57	52	54	41	53	57	34	42	52	61	55	53		
MPNet	54	55	62	64	51	61	59	66	55	65	62	58	41	55	55	50	48	55	58	52	49	45	53	59	29	41	62	62	54	43		
XLM-R-dist	52	51	59	61	45	59	58	64	55	62	61	55	41	53	49	51	47	52	55	52	52	43	54	58	35	46	50	56	49	45		
mT5	50	48	59	56	45	58	56	65	53	63	54	57	39	49	52	39	44	59	52	47	41	39	54	49	35	40	48	52	57	48		
mBERT	46	44	55	53	40	55	49	56	50	49	45	49	36	42	48	44	39	29	47	47	51	37	47	50	30	48	41	54	66	34		
DistilmBERT	44	42	54	50	39	55	45	56	46	50	40	41	35	41	46	40	39	40	45	47	49	36	49	50	26	37	29	54	69	28		
mUSE-dist	50	50	59	58	48	59	54	63	55	60	53	52	42	50	53	47	46	47	59	50	50	37	51	57	31	38	41	57	52	43		
LASER	48	46	55	52	50	55	50	59	54	57	52	52	39	46	46	45	44	44	50	50	48	42	47	52	28	37	43	56	47	38		
mUSE-transformer	45	47	55	55	48	57	52	59	51	56	52	40	43	41	50	42	40	45	46	52	43	39	46	48	28	40	29	54	27	23		
mUSE-cnn	44	45	53	52	44	54	51	57	52	53	51	42	41	42	46	43	38	46	47	47	43	36	49	48	33	48	32	52	27	23		

### BiLSTM Head



## Fine-tuning



## References

- Artetxe, Mikel, and Holger Schwenk. 2019. "Massively Multilingual Sentence Embeddings for Zero-Shot Cross-Lingual Transfer and Beyond." *Transactions of the Association for Computational Linguistics* 7 (September): 597–610. [https://doi.org/10.1162/tacl\\_a\\_00288](https://doi.org/10.1162/tacl_a_00288).
- Augustyniak, Łukasz, Szymon Woźniak, Marcin Gruza, Piotr Gramacki, Krzysztof Rajda, Mikołaj Morzy, and Tomasz Kajdanowicz. 2023. "Massively Multilingual Corpus of Sentiment Datasets and Multi-Faceted Sentiment Classification Benchmark." <https://arxiv.org/abs/2306.07902>.
- Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. "Unsupervised Cross-Lingual Representation Learning at Scale." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440–51. Online: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.747>.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–86. Minneapolis, Minnesota: Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>.
- Feng, Fangxiaoyu, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2020. "Language-agnostic BERT Sentence Embedding." *Computing Research Repository* arXiv:2007.01852. <https://arxiv.org/abs/2007.01852>.
- Rajda, Krzysztof, Łukasz Augustyniak, Piotr Gramacki, Marcin Gruza, Szymon Woźniak, and Tomasz Kajdanowicz. 2022. "Assessment of Massively Multilingual Sentiment Classifiers." In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, 125–40. Dublin, Ireland: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.wassa-1.13>.
- Reimers, Nils, and Iryna Gurevych. 2020. "Making Monolingual Sentence Embeddings Multilingual Using Knowledge Distillation." In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 4512–25. Online: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.emnlp-main.365>.
- Sanh, Victor, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2020. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter." *Computing Research Repository* arXiv:1910.01108. <https://arxiv.org/abs/1910.01108>.
- Xue, Linting, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. "MT5: A Massively Multilingual Pre-Trained Text-to-Text Transformer." In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 483–98. Online: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.naacl-main.41>.
- Yang, Yinfei, Daniel Cer, Amin Ahmad, Mandy Guo, Jax Law, Noah Constant, Gustavo Hernandez Abrego, et al. 2020. "Multilingual Universal Sentence Encoder for Semantic Retrieval." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 87–94. Online: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-demos.12>.