Taxi1500: A Dataset for Multilingual Text Classification in 1500 Languages

Chunlan Ma[1,2], **Ayyoob Imani**[1,2], **Haotian Ye**[1,2], **Renhao Pei**[1],
Ehsaneddin Asgari[3], **Hinrich Schütze**[1,2]
[1]Center for Information and Language Processing (CIS), LMU Munich, Germany
[2]Munich Center for Machine Learning (MCML), Germany
[3]Qatar Computing Research Institute (QCRI), Doha, Qatar
{chunlan, ayyoob, yehao}@cis.lmu.de

Abstract

While broad-coverage multilingual natural language processing tools have been developed, a significant portion of the world's over 7000 languages are still neglected. One reason is the lack of evaluation datasets that cover a diverse range of languages, particularly those that are low-resource or endangered. To address this gap, we present a large-scale text classification dataset encompassing 1504 languages many of which have otherwise limited or no annotated data. This dataset is constructed using parallel translations of the Bible. We develop relevant topics, annotate the English data through crowdsourcing and project these annotations onto other languages via aligned verses. We benchmark a range of existing multilingual models on this dataset. We make our dataset and code available to the public.[1]

1 Introduction

Language inequality is a real issue in the world today as minority languages are under-represented and often excluded from language technologies (Joshi et al., 2020). The lack of technological support for minority languages in communities around the globe has a significant impact on the experience of their users and is commonly a cause for virtual barriers such as the *digital divide*.[2] Recent developments in language technologies have led to a surge in multilingual pre-trained language models (mPLMs), such as mBERT (Devlin et al., 2019), XLM-R (Conneau et al., 2020), and Glot500 (Imani et al., 2023), and large language models (LLMs) like BLOOM (Le Scao et al., 2023) and Aya (Üstün et al., 2024). The lack of knowledge in low-resource languages often causes language technologies to overlook important features from typologically diverse languages (Ponti et al., 2019). A key reason why many low-resource languages remain neglected is the scarcity of evaluation datasets. For example, mPLMs like mBERT and XLM-R are evaluated on much fewer languages than they are trained for, largely due to the limited availability of languages in most existing benchmark datasets.

As a solution, we propose a dataset that covers more than 1500 languages. We use translations of the Bible as our source and develop topics that are well generalized (so as to apply to many verses), but at the same time are not overly abstract. We obtain annotations for the English verses using crowdsourcing. Because the Bible is aligned at the verse level, we can easily project annotations from the English side to all other languages. To ensure the quality of our annotated data, we calculate the inter-annotator agreement using Krippendorff's α. In addition, we introduce a benchmark for four mPLMs and three LLMs. We present evaluation results using mBERT, XLM-R-Base, XLM-R-Large and Glot500 for all languages and LLaMA2-7B (Touvron et al., 2023), Mistral-7B (Jiang et al., 2023), and BLOOM (560m, 1B, 3B and 7B) for 64 selected languages in our dataset. Glot500 demonstrates better multilingual capabilities, attributed to its larger number of languages in the pretraining data. Moreover, the evaluation of LLMs reveals that their performance (based on a few low-resource prompts) is comparable to fine-tuned mPLMs.

2 Related Works

To date, most datasets that can be used for multilingual task evaluation (Pan et al., 2017; Conneau et al., 2018; De Marneffe et al., 2021; Adelani et al., 2021, 2024; Adebara et al., 2022) cover no more than a few hundred languages, a small number compared to the world's 7000 languages. In current NLP research, parallel corpora play a crucial role as they serve as cross-lingual bridges, enabling the processing and understanding of less known languages through other languages. In this study, we employ translations of the Bible as the source of

[1]https://github.com/cisnlp/Taxi1500
[2]labs.theguardian.com/digital-language-divide

Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers), pages 414–439
April 29 - May 4, 2025

parallel data, utilizing both the Parallel Bible corpus (Mayer and Cysouw, 2014), covering 1304 languages, as well as 1000Langs,[3] Bible translations collected from multiple Bible websites, resulting in a total coverage of 1504 languages.

3 Dataset Creation

Since many low-resource languages only have a translated New Testament, we use verses from the New Testament to build our dataset. In the initial annotation phase, we gather topics using Latent Dirichlet Allocation (LDA),[4] online preaching websites with topics of Bible verses,[5] and insights from linguists. We then utilize Amazon Mechanical Turk (MTurk)[6] for crowdsourcing to assess the quality of the selected topics. We conduct seven rounds of topic selection and show the details in Table 7 in Appendix E. Ultimately, we choose the six topics with the most verses: *recommendation*, *faith*, *description*, *sin*, *grace*, and *violence*. Following this, three annotators extract verses for each of the six topics, selecting only those where at least two annotators agree. We remove verses that cover multiple topics or are not relevant to any topic as such noise complicates annotation and may confuse crowdsourcing annotators. This curation reduces annotation cost. We then submit the resulting 1,077 verses to Amazon MTurk, specifying the US as the annotators' location. Each verse is annotated ten times, with final labels determined by majority voting.

We assume annotation quality issues may arise if 1) the task is confusing, or 2) the worker lacks care or attention. We provide detailed guidelines and examples along with the task. All workers must also pass a qualification test to ensure they fully understand the task. For quality control, we implement a performance threshold. We create "pseudo gold standard" data based on majority votes from all annotators and calculate each worker's macro F1 score. If that score is below 0.40 for a worker, their annotations are rejected, and the verses are republished for re-annotation.

We use Krippendorff's α (K-α) to compute inter-annotator agreement. K-α is chosen for its ability to handle missing annotations in the dataset. This is important because each worker only annotates

[3]https://github.com/ehsanasgari/1000Langs
[4]https://tinyurl.com/5fja5yvz
[5]https://www.georgeho.org/lda-sucks/
[6]www.mturk.com

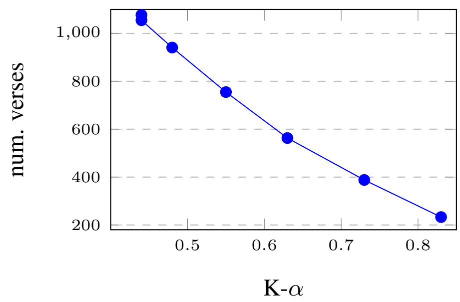

Figure 1: Tradeoff between K-α and the number of verses. Each dot in the plot stands for a threshold of the required minimum votes $\in \{3, 4, 5, 6, 7, 8, 9\}$ for a verse to be accepted.

a subset of the verses. Table 2 shows K-α values for different thresholds, i.e., the minimum votes for the majority label required for a verse to be accepted. We obtain K-α = 0.44 on the entire dataset, which can be improved by raising the threshold of required votes. But as Figure 1 demonstrates, there is a clear tradeoff between the number of accepted verses and K-α, and increasing K-α reduces the size of the dataset significantly. Furthermore, a slightly suboptimal K-α value is predictable considering that the topics of our task are rather subjective due to the highly specialized domain. Also, as (Price et al., 2020) points out, a low K-α does not necessarily signify low data quality. We thus do not remove any data by raising the required number of votes but instead rely on our control measures (e.g., removing annotations by unreliable crowdworkers) to ensure data quality.

4 Dataset

The final dataset we obtain consists of verses categorized into six topics: *faith*, *grace*, *sin*, *violence*, *description*, and *recommendation*. Table 1 shows an overview of the topics with one example for each, as well as the number of verses of each topic in the English dataset. Class *violence*, with 59 instances, is the smallest class and *recommendation*, with 281, is the biggest class. Since some languages have incomplete translations of the New Testament that do not contain all of the 1077 verses, we exclude languages where the total number of annotated verses is fewer than 900. This leaves us with 1504 languages from 113 language families which are spread across the globe.[7]

[7]Family and geographical data from glottolog.org

class	example	num. verses
recommendation	If you love me, you will observe my commandments	281
faith	Most truly I say to you, whoever believes has everlasting life	260
description	There was a man of the Pharisees named Nicodemus, a ruler of the Jews	184
sin	Jesus answered: “I do not have a demon, but I honor my Father, and you dishonor me	153
grace	The Father loves the Son and has given all things into his hand	140
violence	He put James the brother of John to death by the sword	59

Table 1: An overview of the six classes of our dataset, with one example verse and the number of verses in the crowdsourced English dataset for each class.

vote $\geq$	3	4	5	6	7	8	9
num. verses	1077	1055	941	755	563	388	233
K-α	0.44	0.44	0.48	0.55	0.63	0.73	0.83

Table 2: The K-α value increases as we specify a higher threshold for the minimum number of votes of the majority topic.

5 Benchmarking

To illustrate its utility, we use Taxi1500 to evaluate four pre-trained multilingual models: mBERT, XLM-R-Base, XLM-R-Large, and Glot500, and three LLMs: LLaMA2, BLOOM, and Mistral using a selection of 64 languages from Taxi1500. For a fair comparison, we split languages in our dataset into three subsets, namely head languages, Glot500-only languages, and tail languages. Head languages are languages that are in the pre-training data of all four models. Glot500-only languages are languages that are only in the pre-training data of Glot500. Tail languages are languages that are not in the pre-training data of any model. Details of the setup are provided in Appendix A.

5.1 Experiment Setup

Our experiments are divided into three settings: zero-shot transfer, in-language classification, and three-shot prompting for LLMs. The dataset for each of the 1,504 languages is split into training, development, and test sets with an 80/10/10 ratio.

In the in-language classification setting, we use the target language data for fine-tuning and testing. In zero-shot transfer, we use English data for fine-tuning and test on the target language test set. For in-language experiments on languages other than English, we furthermore vary the training set size $\in \{50, 100, 200, 400, 600, 860\}$, where 860 corresponds to the full training set, in order to test: 1) the effects of different amounts of training samples and 2) the minimal number of training samples required to achieve acceptable classification results.

5.2 Results

Zero-shot transfer. We conduct Bag-of-Words (BOW) classification with our dataset as a baseline and present the results in Appendix I. The results revealed extremely low accuracy for BOW: most of the results are less than 0.10, indicating that to classify verses in our dataset correctly, the models must have access to a good semantic representation (which BOW does not seem to provide).

In Figure 2, we show the results for 1504 languages, divided into three sets: head languages (left), Glot500-only languages (middle), and tail languages (right). On head languages, Glot500, XLM-R-B, and XLM-R-L have 68, 65, and 69 languages within the high F1 range (0.4-0.8), respectively, while mBERT only has 26 languages within this range, indicating its worse performance. This might be explained by a smaller amount of pre-training data of mBERT compared with the other three models. On Glot500-only languages, Glot500 outperforms the other three models with 117 languages in the range of 0.2-0.8, whereas the other three models have fewer than 30 languages within this range. Because Glot500-only languages are in the pre-training data of Glot500, we expect Glot500 to achieve better results on these languages. On tail languages, Glot500 outperforms the other three models slightly with around 100 fewer languages in the range of 0-0.2. The reason might be that a larger number of pre-training languages contributes to higher performance for other tail languages from the same family. The zero-shot transfer results indicate that Taxi1500 can effectively demonstrate better performance for models pretrained using more languages.

In-language training. To investigate the influence of the training set size, we conduct in-language experiments with 20 languages (10 head and 10 tail languages), which are selected to repre-

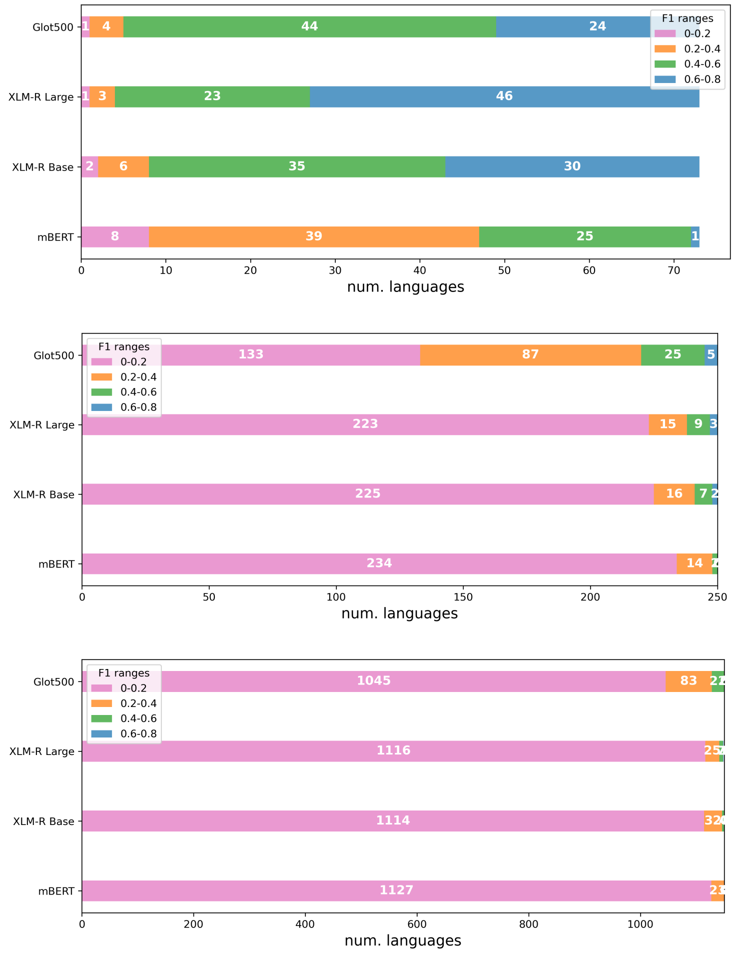

Figure 2: Zero shot transfer learning: head languages (left), Glot500-only languages (middle), and tail languages (right). X-axis is the number of languages, y-axis presents four models. We split F1 scores into four ranges: 0-0.2, 0.2-0.4, 0.4-0.6 and 0.6-0.8.

head lang.	zero shot	in-language training 50	100	200	400	600	860	tail lang.	zero shot	in-language training 50	100	200	400	600	860
eng	0.65	0.35	0.33	0.53	0.49	0.51	0.71	chr	0.09	0.15	0.20	0.15	0.24	0.21	0.28
deu	0.52	0.16	0.18	0.43	0.49	0.52	0.51	gag	0.33	0.17	0.13	0.14	0.45	0.32	0.54
heb	0.15	0.10	0.13	0.18	0.16	0.33	0.35	hix	0.06	0.18	0.17	0.22	0.3	0.43	0.49
jpn	0.62	0.25	0.39	0.53	0.57	0.61	0.68	hlt	0.05	0.14	0.07	0.19	0.40	0.20	0.50
kaz	0.57	0.23	0.35	0.47	0.41	0.55	0.56	kpv	0.09	0.09	0.21	0.23	0.41	0.38	0.53
kor	0.63	0.35	0.55	0.58	0.65	0.53	0.70	kum	0.13	0.13	0.17	0.22	0.27	0.37	0.45
eus	0.26	0.09	0.26	0.25	0.34	0.37	0.34	luc	0.11	0.12	0.11	0.30	0.30	0.39	0.39
mal	0.42	0.18	0.30	0.21	0.45	0.45	0.64	mag	0.38	0.11	0.23	0.41	0.48	0.38	0.51
pes	0.66	0.17	0.55	0.47	0.65	0.64	0.71	mbd	0.11	0.18	0.14	0.25	0.30	0.30	0.38
zho	0.63	0.33	0.49	0.52	0.45	0.51	0.68	npl	0.05	0.14	0.08	0.25	0.41	0.41	0.43
avg.	0.51	0.22	0.35	0.42	0.47	0.50	0.59	**avg.**	0.14	0.14	0.15	0.24	0.36	0.34	0.45

Table 3: Results of zero-shot transfer and in-language fine-tuning experiments using XLM-R-Base for 20 selected languages, 10 head (left): English, German, Hebrew, Japanese, Kazakh, Korean, Basque, Malayalam, Persian and Chinese, and 10 tail (right): Cherokee, Gagauz, Hixkaryana, Nga La, Komi-Zyrian, Kumyk, Aringa, Magahi, Dibabawon Manobo and Southeastern Puebla Nahuatl. The numbers in the table header indicate the size of target language training data: 860 means the full training set.

sent a diverse range of languages from 13 families. Tables 3 and 4 show the results of zero-shot transfer and in-language experiments using mBERT and XLM-R-B for the selected languages. As expected,

head lang.	zero shot	in-language training 50	100	200	400	600	860	tail lang.	zero shot	in-language training 50	100	200	400	600	860
eng	0.71	0.35	0.33	0.53	0.49	0.51	0.71	chr	0.05	0.24	0.21	0.29	0.35	0.30	0.35
deu	0.39	0.20	0.13	0.34	0.42	0.44	0.52	gag	0.12	0.21	0.29	0.35	0.39	0.45	0.38
heb	0.36	0.24	0.24	0.36	0.33	0.38	0.41	hix	0.07	0.30	0.27	0.35	0.35	0.39	0.41
jpn	0.39	0.37	0.40	0.32	0.49	0.63	0.66	hlt	0.08	0.16	0.25	0.33	0.34	0.44	0.49
kaz	0.29	0.30	0.36	0.38	0.50	0.48	0.48	kpv	0.08	0.19	0.24	0.45	0.41	0.39	0.46
kor	0.41	0.36	0.36	0.45	0.56	0.50	0.60	kum	0.14	0.28	0.27	0.35	0.37	0.42	0.46
eus	0.17	0.15	0.12	0.31	0.44	0.46	0.43	luc	0.08	0.27	0.23	0.46	0.41	0.45	0.35
mal	0.22	0.32	0.31	0.41	0.41	0.40	0.46	mag	0.19	0.14	0.38	0.38	0.37	0.43	0.34
pes	0.43	0.30	0.36	0.55	0.53	0.52	0.56	mbd	0.08	0.18	0.33	0.36	0.36	0.39	0.42
zho	0.36	0.24	0.46	0.47	0.62	0.54	0.59	npl	0.06	0.21	0.30	0.38	0.39	0.40	0.40
avg.	0.37	0.28	0.31	0.41	0.48	0.49	0.54	**avg.**	0.10	0.22	0.28	0.37	0.37	0.41	0.41

Table 4: Results of zero-shot transfer and in-language fine-tuning experiments using mBERT for 20 selected languages, 10 head (left): English, German, Hebrew, Japanese, Kazakh, Korean, Basque, Malayalam, Persian and Chinese, and 10 tail (right): Cherokee, Gagauz, Hixkaryana, Nga La, Komi-Zyrian, Kumyk, Aringa, Magahi, Dibabawon Manobo and Southeastern Puebla Nahuatl. The numbers in the table header indicate the size of target language training data: 860 means the full training set.

Model	LLaMA2	Mistral	BLOOM			
Size	7B	7B	560M	1B	3B	7B
Avg. Acc	0.45	0.55	0.46	0.50	0.48	0.48

Table 5: Performance of three LLMs of various sizes.

the in-language performance improves when the training set becomes larger. Interestingly, zero-shot transfer performance of head languages is comparable to in-language setting with 100 samples for mBERT and with 400 samples for XLM-R-B, which indicates that models with more parameters may require more in-language data to reach a comparable level with zero-shot transfer performance. Moreover, the zero-shot transfer results on both models show that head languages consistently outperform tail languages, which reflects both models' better generalization capability on languages in their pretraining data.

Evaluation of LLMs. To explore the capability of LLMs, we conduct three-shot in-context learning with 64 selected languages from different language families on six LLMs, namely LLaMA2-7B, Mistral-7B, and BLOOM (560m, 1B, 3B and 7B). We report the results in Appendix H. In Table 5, we show the average score of 64 languages. Notably, Mistral-7B achieves the highest average performance with a score of 0.55, surpassing both LLaMA2-7B, which scores 0.45, and BLOOM at various sizes. BLOOM's performance varies slightly across model sizes, with the 1B version yielding the highest score (0.50) among BLOOM models, while the 7B version underperforms at 0.46. These results suggest that Mistral-7B may be more effective in handling the Taxi1500 task. Overall, each LLM achieves performance comparable to the mPLMs on in-language classification tasks trained on a full training set of 860 verses. This result could be interpreted as LLMs having multilingual capabilities similar to mPLMs (even though the LLM setup requires no finetuning training data). But of course this experiment was only conducted on 64 languages. It remains to be verified that it generalizes to low-resource languages in general.

6 Conclusion

In this paper, we propose a text classification dataset consisting of 1504 languages by annotating English Bible verses through crowdsoucing and projecting the labels to other languages with parallel data. We benchmark several widely used multilingual language models and LLMs using our dataset. The results demonstrate that Taxi1500 can effectively evaluate multilingual capabilities across different models.

7 Limitations

While the high degree of parallelism in the PBC makes it a valuable tool for massively multilingual application, such as the building of our evaluation dataset, it is not perfect. One limitation is the specific domain of the Bible being a religious text, which often does not reflect real world usages. The specific religious context additionally makes it possible that keywords are exploited. Also, we are restricted to the New Testament as a large quantity of languages do not have a translated Old Testament in the PBC. Given that some extremely low-resource languages do not have complete translations, the actual number of available verses varies for each language. However, since the Bible is by far the most translated book in the world, we regard it as a suitable resource for an initiative to build highly parallel data like ours.

8 Ethics Statement

In this work, we introduce a new multilingual text classification dataset based on the Parallel Bible Corpus. The data is partially annotated by workers from the Amazon mTurk platform, who are rewarded fairly for their work ($0.2 per sentence). Our dataset contains Bible verses for which we estimate a low risk of tracing to specific individuals and are intended exclusively for the evaluation of NLP tasks concerning the supported languages. We therefore do not expect any ethical issues with our dataset.

Bird (2024) has argued that many low-resource languages (in particular, languages that are primarily used orally) do not benefit from NLP technology and may even be harmed, e.g., if social media companies' use of low-resource NLP technology results in younger speakers of a low-resource language spending more time on their devices and less time engaging with their community. We acknowledge that this is a real danger for some low-resource communities. We also believe that the benefits of NLP outweigh the risks for others, e.g., for Occitan. In general, this is an important question about the future direction of NLP research that goes beyond this paper.

9 Acknowledgements

This work was funded by the European Research Council (NonSequeToR, grant #740516) and by DFG (SCHU 2246/14-1). We appreciate Yihong Liu's suggestions for the revisions during the writing process of the paper. We also thank the anonymous reviewers for their constructive feedback.

References

Ife Adebara, AbdelRahim Elmadany, Muhammad Abdul-Mageed, and Alcides Inciarte. 2022. AfroLID: A neural language identification tool for African languages. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 1958–1981, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

David Adelani, Hannah Liu, Xiaoyu Shen, Nikita Vassilyev, Jesujoba Alabi, Yanke Mao, Haonan Gao, and En-Shiun Lee. 2024. SIB-200: A simple, inclusive, and big evaluation dataset for topic classification in 200+ languages and dialects. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 226–245, St. Julian's, Malta. Association for Computational Linguistics.

David Ifeoluwa Adelani, Jade Abbott, Graham Neubig, Daniel D'souza, Julia Kreutzer, Constantine Lignos, Chester Palen-Michel, Happy Buzaaba, Shruti Rijhwani, Sebastian Ruder, Stephen Mayhew, Israel Abebe Azime, Shamsuddeen H. Muhammad, Chris Chinenye Emezue, Joyce Nakatumba-Nabende, Perez Ogayo, Aremu Anuoluwapo, Catherine Gitau, Derguene Mbaye, Jesujoba Alabi, Seid Muhie Yimam, Tajuddeen Rabiu Gwadabe, Ignatius Ezeani, Rubungo Andre Niyongabo, Jonathan Mukiibi, Verrah Otiende, Iroro Orife, Davis David, Samba Ngom, Tosin Adewumi, Paul Rayson, Mofetoluwa Adeyemi, Gerald Muriuki, Emmanuel Anebi, Chiamaka Chukwuneke, Nkiruka Odu, Eric Peter Wairagala, Samuel Oyerinde, Clemencia Siro, Tobius Saul Bateesa, Temilola Oloyede, Yvonne Wambui, Victor Akinode, Deborah Nabagereka, Maurice Katusiime, Ayodele Awokoya, Mouhamadane MBOUP, Dibora Gebreyohannes, Henok Tilaye, Kelechi Nwaike, Degaga Wolde, Abdoulaye Faye, Blessing Sibanda, Orevaoghene Ahia, Bonaventure F. P. Dossou, Kelechi Ogueji, Thierno Ibrahima DIOP, Abdoulaye Diallo, Adewale Akinfaderin, Tendai Marengereke, and Salomey Osei. 2021. MasakhaNER: Named entity recognition for African languages. *Transactions of the Association for Computational Linguistics*, 9:1116–1131.

Steven Bird. 2024. Must NLP be extractive? In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14915–14929, Bangkok, Thailand. Association for Computational Linguistics.

Alexis Conneau, 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*, pages 8440–8451, Online. Association for Computational Linguistics.

Alexis Conneau, Ruty Rinott, Guillaume Lample, Adina Williams, Samuel Bowman, Holger Schwenk, and Veselin Stoyanov. 2018. XNLI: Evaluating cross-lingual sentence representations. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2475–2485, Brussels, Belgium. Association for Computational Linguistics.

Marie-Catherine De Marneffe, Christopher D Manning, Joakim Nivre, and Daniel Zeman. 2021. Universal dependencies. *Computational linguistics*, 47(2):255–308.

Jacob Devlin, 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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Philipp Dufter, Mengjie Zhao, Martin Schmitt, Alexander Fraser, and Hinrich Schütze. 2018. Embedding learning through multilingual concept induction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1520–1530, Melbourne, Australia. Association for Computational Linguistics.

Ayyoob Imani, Peiqin Lin, Amir Hossein Kargaran, Silvia Severini, Masoud Jalili Sabet, Nora Kassner, Chunlan Ma, Helmut Schmid, André Martins, François Yvon, and Hinrich Schütze. 2023. Glot500: Scaling multilingual corpora and language models to 500 languages. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1082–1117, Toronto, Canada. Association for Computational Linguistics.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b.

Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online. Association for Computational Linguistics.

Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilić, Daniel Hesslow, Roman Castagné, Alexandra Sasha Luccioni, François Yvon, Matthias Gallé, et al. 2023. Bloom: A 176b-parameter open-access multilingual language model.

Peiqin Lin, Shaoxiong Ji, Jörg Tiedemann, André F. T. Martins, and Hinrich Schütze. 2024. Mala-500: Massive language adaptation of large language models.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach.

Thomas Mayer and Michael Cysouw. 2014. Creating a massively parallel Bible corpus. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 3158–3163, Reykjavik, Iceland. European Language Resources Association (ELRA).

Xiaoman Pan, Boliang Zhang, Jonathan May, Joel Nothman, Kevin Knight, and Heng Ji. 2017. Cross-lingual name tagging and linking for 282 languages. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1946–1958, Vancouver, Canada. Association for Computational Linguistics.

Edoardo Maria Ponti, Helen O'Horan, Yevgeni Berzak, Ivan Vulić, Roi Reichart, Thierry Poibeau, Ekaterina Shutova, and Anna Korhonen. 2019. Modeling language variation and universals: A survey on typological linguistics for natural language processing. *Computational Linguistics*, 45(3):559–601.

Ilan Price, Jordan Gifford-Moore, Jory Flemming, Saul Musker, Maayan Roichman, Guillaume Sylvain, Nithum Thain, Lucas Dixon, and Jeffrey Sorensen. 2020. Six attributes of unhealthy conversations. In *Proceedings of the Fourth Workshop on Online Abuse and Harms*, pages 114–124, Online. Association for Computational Linguistics.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models.

Ahmet Üstün, Viraat Aryabumi, Zheng Yong, Wei-Yin Ko, Daniel D'souza, Gbemileke Onilude, Neel Bhandari, Shivalika Singh, Hui-Lee Ooi, Amr Kayid, Freddie Vargus, Phil Blunsom, Shayne Longpre, Niklas Muennighoff, Marzieh Fadaee, Julia Kreutzer, and Sara Hooker. 2024. Aya model: An instruction fine-tuned open-access multilingual language model. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15894–15939, Bangkok, Thailand. Association for Computational Linguistics.

A Experiment setup

We fine-tune four mPLMs using the training data for each setting. We use the AdamW optimizer with a learning rate of $2e-5$ and a batch size $\in \{16, 32\}$, and report the best metrics. Training is stopped employing early stopping on the development data. All experiments can be performed on a single GeForce GTX 1080Ti GPU within a matter of minutes.

B In language training results

In the in-language classification setting, we use the target language data for fine-tuning and testing. In zero-shot transfer, we use English data for fine-tuning and get predictions on the target language test set. For in-language experiments on languages other than English, we furthermore vary the training set size $\in \{50, 100, 200, 400, 600, 860\}$, where 860 corresponds to the full training set.

C Analysis by Language Family

In Figures 3 and 4, we present zero-shot transfer and in-language results of all languages based on their families on XLM-R-Base and Glot500. For almost all families, the performance on head languages is significantly higher than that of Glot500-only and tail languages. The Indo-European family outperforms other language families not only on head languages but also on Glot500-only and tail languages. We suppose the reason is that the four evaluated models are pre-trained with more Indo-European languages, which increases the performance of this family. We also notice that XLM-R-Large tends to perform worse than the other three models on most languages. We think this could be due to its larger number of parameters, which makes it prone to overfitting on our small dataset. Interestingly, by comparing zero-shot transfer and in-language results of XLM-R-Base, we find that languages that are extremely low-resource and use non-Latin scripts (e.g. Yawa-Saweru, Lengua-Mascoy, and Hmong-Mien) have significant performance increases (around 0.4) when they are trained with in-language data. This indicates that the four models do not perform as well on non-Latin scripts as on Latin scripts.

D Annotation

Figure 5 shows a screenshot of the annotation interface. Workers select one label for each verse among six options. If they think one verse does not belong to any of them, the workers should classify this verse as *Other*.

E Topics Design

We present our attempts to explore the classification task and the construction of possible categories. There are different classification tasks, for example, sentiment classification, intent classification, and topic classification. At the beginning, we attempt to implement sentiment classification and split verses into three conventional categories: positive, neutral, and negative. However, most of the verses in the Bible do not indicate one absolute sentiment. Hence, we try intent classification yet also failed. We demonstrate this in more detail below.

E.1 Sentiment Classification

First, we attempt to implement the simplest sentiment classification task. Dufter et al. (2018) classify a portion of the English verses in the PBC into a positive category and a negative category. Inspired by them, we initially try standard sentiment classification on the PBC with an improved method from Dufter et al. (2018). Precisely, in order to explore the possibility of using more categories, we divide verses in the Bible into positive, negative and neutral ones using the prepared sentiment RoBERTa model (Liu et al., 2019) from Huggingface, which is fine-tuned on 5,304 manually annotated social media posts with 86.1% accuracy. We get 6,233 negative verses, 1,441 negative verses, and 23,459 neutral verses from a total of 31,133 verses from `eng-x-bible-newworld2013.txt` (considering the entire Bible, rather than only the New Testament, which results in a much higher verse count).

We propose to conduct emotion classification on positive and negative verses because we assume these verses have a higher probability of containing emotions. We utilize a fine-tuned DistilBERT model[8] to perform emotion classification, which is a multi-class classification task with six labels: Joy, Anger, sadness, Fear, Love, and Surprise. The numbers of verses in each category are as follows: Sadness: 1171, Joy: 1952, Love: 870, Anger: 4201, Fear: 457, Surprise: 29. However, a great num-

[8] https://huggingface.co/bhadresh-savani/distilbert-base-uncased-emotion

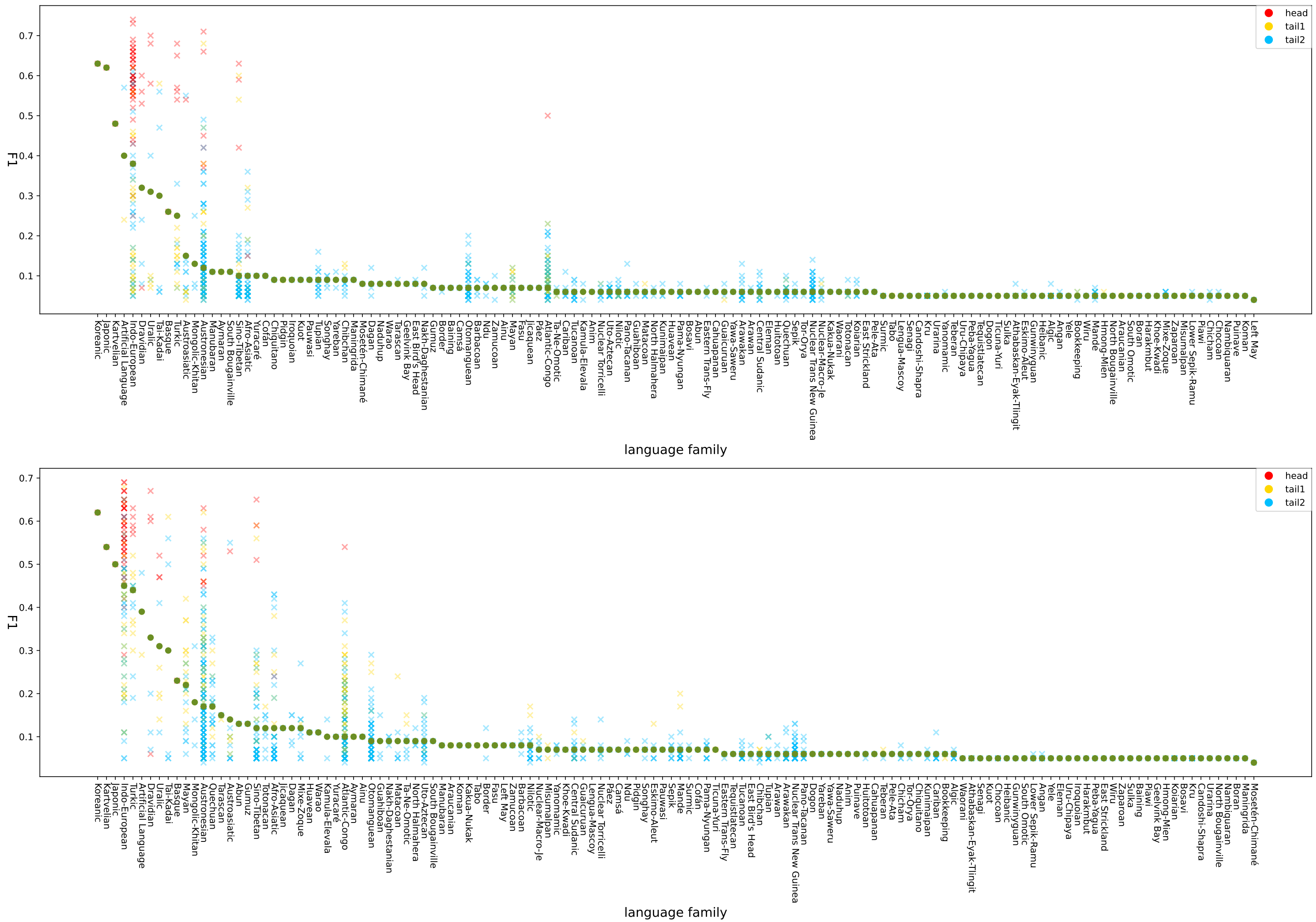

Figure 3: Zero shot transfer learning: F1 of XLM-R-Base (top) and Glot500 (bottom). Each small dot represents a language, each large dot an average per family. Families are sorted by F1. Red, yellow and blue represent head, Glot500-only and tail languages respectively.

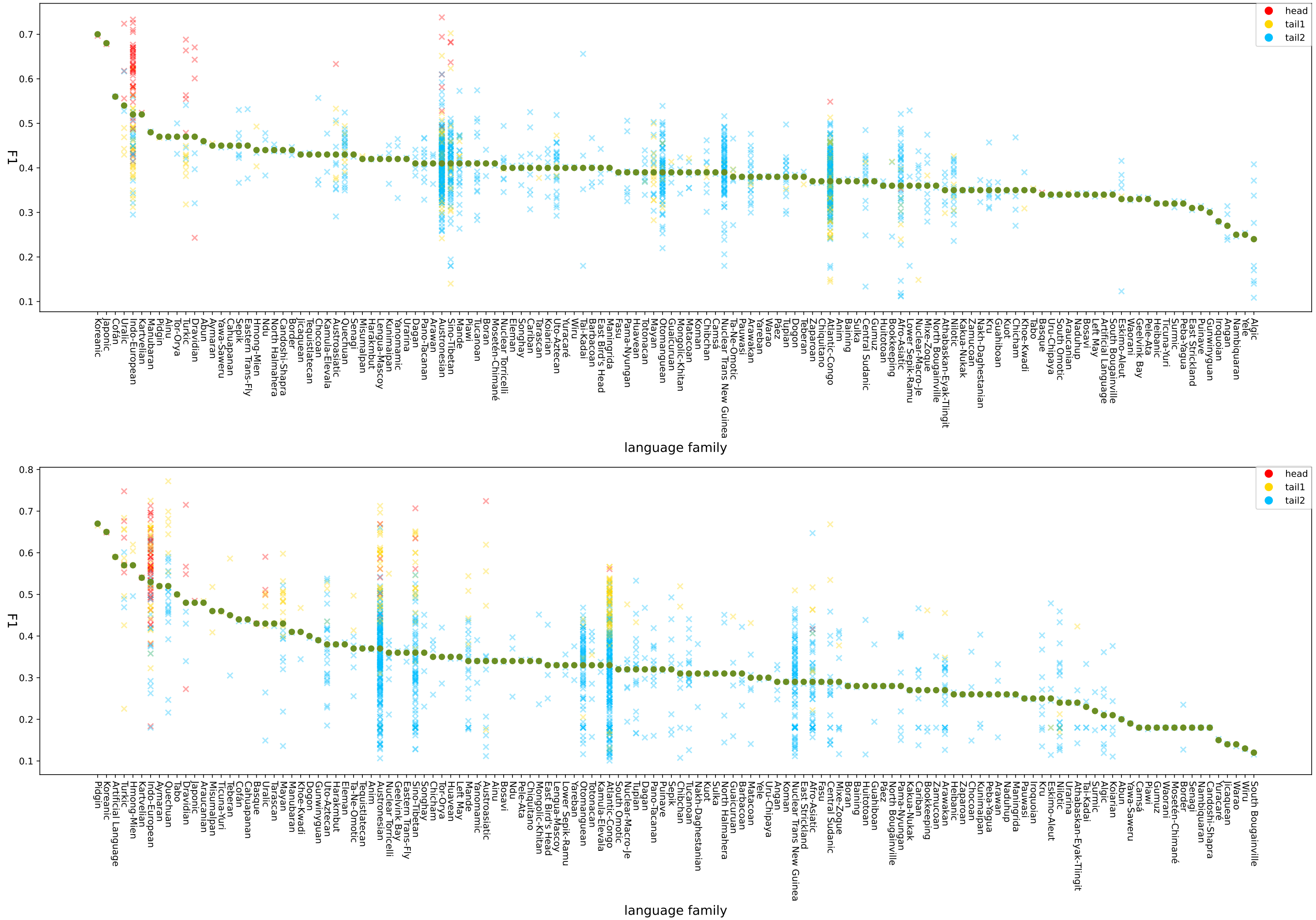

Figure 4: In-language results: F1 of XLM-R-Base (top) and Glot500 (bottom). Each small dot represents a language, each large dot an average per family. Families are sorted by F1. Red, yellow and blue represent head, Glot500-only and tail languages respectively.

head lang.	iso	Script	Family	tail lang.	iso	Script	Family
English	eng	Latin	Indo-European	Cherokee	chr	Cherokee	Iroquoian
German	deu	Latin	Indo-European	Gagauz	gag	Latin	Turkic
Hebrew	heb	Hebrew	Afro-Asiatic	Hixkaryana	hix	Latin	Cariban
Japanese	jpn	Japanese	Japanic	Nga La	hlt	Latin	Sino-Tibetan
Kazakh	kaz	Cyrilic	Turkic	Komi-Zyrian	kpv	Cyrilic	Uralic
Korean	kor	Korean	Koreanic	Kumyk	kum	Cyrilic	Turkic
Basque	eus	Latin	Basque	Aringa	luc	Latin	Central Sudanic
Malayalam	mal	Malayalam	Dravidian	Magahi	mag	Devanagari	Indo-European
Persian	pes	Arabic	Indo-European	Dibabawon Manobo	mbd	Latin	Austronesian
Chinese	zho	Chinese	Sino-Tebietan	Middle Watut	npl	Latin	Uto-Aztecan

Table 6: An overview of selected 20 languages from 11 different writing systems and 13 language families

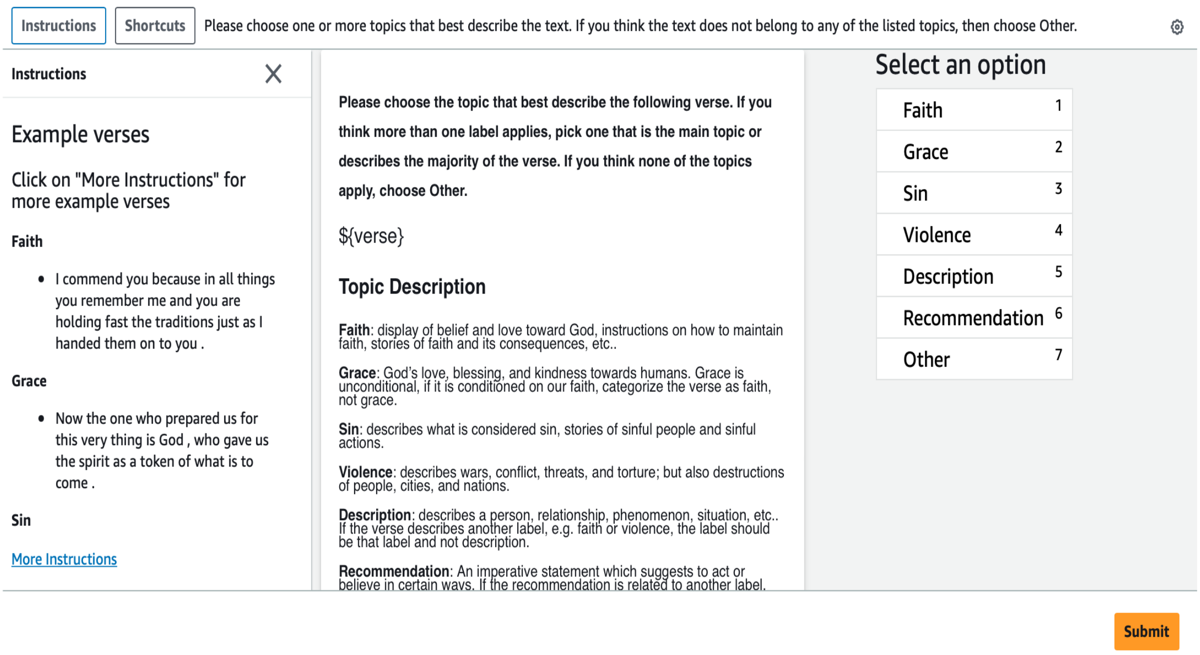

Figure 5: mTurk interface with English instructions and verse examples

ber of verses are not correctly classified because most verses in the Bible are objective, and it is impossible to classify them into emotions. For example, verse `01037029 Later when Reuben returned to the waterpit and saw that Joseph was not in the waterpit, he ripped his garments apart` is an objective sentence, but is assigned an emotion *Anger*. Thus, we do not use the emotion classification task and instead continue to seek other categories.

E.2 Category Design

The failure of emotion classification implies that the Bible verses are not suitable for subjective classification. We thus decide to design topic categories using *Latent Dirichlet Allocation* topic search model.[9]

[9]https://tinyurl.com/mr487nc6

E.2.1 Latent Dirichlet Allocation

To detect latent topics in the Bible, We use the Latent Dirichlet Allocation topic model. We set the number of tokens to describe each topic to 10 and the number of topics to 200. Besides eliminating the common stop words with `NLTK stopwords` package, we also filter out highly frequent words such as *God* and *Jehova*, and meaningless tokens like *ah* and *el*. However, LDA produces results that do not indicate meaningful topics based on the output words. We present five randomly chosen sets of words to show an example:

Topic 1: [house, people, one, may, david, sons, become, day, according, saying]
Topic 2: [david, son, one, house, man, things, came, king, hand, land]
Topic 3: [sons, israel, one, like, king, house, man,

people, us, men]
Topic 4: [land, one, let, people, men, us, went, took, go, brought]
Topic 5: [one, israel, king, people, may, like, man, days, seven, moses]

We can observe that there are many overlapping words in different topics, and it is difficult to interpret the results. The reason is presumably that LDA is suitable for processing long documents, but a verse normally contains fewer than 50 tokens and is too short to extract hidden topics for LDA. In addition, LDA may not work well on documents that do not coherently discuss a single topic, and there are numerous verses that do not belong to just one specific topic. A Reddit comments classification experiment by other researcher also occurs the same problem as ours.[10]

E.2.2 Self-Designed Categories

Because LDA fails to produce meaningful topics of the Bible verses, we attempt to create some categories according to commonly occurring verses, v1 in table 7 shows the initial category design. The first version contains categories *Rules, Phenomenon, Conflict, Relation, Place, Character, Reward, Punishment, and Command*. *Rules* defines verses that state an activity must or must not be done. *Phenomenon* describes natural or societal facts. *Conflict* includes argument, violence, or war among people, groups, or countries. *Relation* reflects family genealogy. *Place* includes verses that contain a city or area where an event happens. *Character* contains verses that indicate the personality of a person. *Reward* describes a person given something by God because he has done something good. *Punishment* is the counterpart of reward that describes punishment from God. *Command* is the order from God. After the categories are defined, we look for several example verses that can be shown to crowdsource workers in order to annotate the data. However, by collecting example verses, we find overlapping definitions between certain categories. For instance, the verse `03019023 When you come into the land and you plant any tree for food , you must consider its fruitage impure and forbidden . For three years it will be forbidden to you . It must not be eaten .` can be either annotated as *Command* or *Rules*. Therefore, in order to obtain better categories and alleviate the category overlap, we seek help from topic models and experts. The next paragraphs present details on exploring the categories.

E.2.3 Online Bible Topics

Following the failure of self-designed categories, we analyze the difficulty to create categories for the Bible verses. Compared with data of other benchmarks that normally use Common Crawl or Wikipedia, the domain of the Bible is too specific to extract categories merely according to common sense. Instead, theological knowledge may assist in category creation. Thus, we change the strategy of building categories by browsing websites with the keywords "*Bible topics*". Thanks to a large number of available preaching websites, we are able to find a lot of topics created to help with the creation of categories. Those topics are presented on the websites with verses examples. Among all websites we have browsed, ProPreacher[12] is the best one with a variety of 100 sermon topics and respective verse examples. Subsequently, we select topics from 100 sermon topics. There are two principles when selecting topics. First, we ensure that the benchmark is challenging, thus more categories should be contained. Second, in order to build a dataset with enough sentences, only topics with many examples should be chosen. In the end, we collect 15 categories (v2 in table 7) with sufficient example verses. Before we start crowsourcing with these categories, we show three NLP students the category collection and 100 randomly sampled verses to annotate. They reflect that these topics are too abstract to understand. For example, *Eschatology*, *Philosophy* and *Theology* are hard to apply to respective verses. Therefore, we adjust the categories to v3 (table 7) based on v2 and the feedback. v3 deletes abstract topics *Eschatology, Philosophy, Theology, and Moral*, while adding *Repentance, Friendship, Thankfulness, Forgiveness, and Suffering* that are collected from other preaching websites. The topic *Persecution* is changed to *Heresy*. Once finished the task and category design, we start with crowdsourcing to obtain annotated data.

[10] `https://www.georgeho.org/lda-sucks/`

[12] `https://www.propreacher.com/100-sermon-topics/`

version	Category	Num of category
v1	Rules, Phenomenon, Conflict, Relation, Place, Character, Reward, Punishment, Command	9
v2	Eschatology, Grace, Family, Creation, Philosophy, Revival, Cults, Compromise, Persecution, Hospitality, Conflicts, Theology, Morals, Commandments, Sacrifice	15
v3	Creation, Grace, Violence, Conflict, Hospitality, Sacrifice, Heresy, Repentance, Faith, Suffering, Forgiveness, Thankfulness, Friendship, Temptation	14
v4	Creation, Grace, Violence, Conflict, Hospitality, Sacrifice, Heresy, Repentance, Faith, Suffering, Forgiveness, Thankfulness	12
v5	Creation, Commandment, Genealogy, Violence, Sacrifice, Money, Salvation, Sin	8
v6	Creation, Commandment, Genealogy, Violence, Sacrifice, Money, Grace, Sin	8
v7	Recommendation, Faith, Description, Sin, Grace, Violence	6

Table 7: Different versions of designed categories. v1 is the initial self-designed version with the help of a linguist. v2 is collected based on online preaching websites ProPreacher[11]. v3 deletes three abstract labels *Eschatology, Philosophy, Theology*, and *Moral*, and adds four new labels *Repentance, Friendship, Thankfulness, Forgiveness* and *Suffering*. v4 is the version we use to crowdsource annotation on Amazon Mechanical Turk. v5 and v6 combines similar labels of v4 and changes the names of several labels. v7 is the version we use for our final dataset.

E.2.4 Crowdsource Attempts

We choose Amazon Mechanical Turk (mTruk) to test the designed topics because of its availability of a large number of native English speakers that we are looking for. Besides, it has sufficient online tutorials that can help to build annotation projects. When the v3 (table 7) class design is determined, we use mTurk to assign verses and test the quality of designed topics.

F Data collection

Our dataset is built based on PBC and 1000Langs. Due to the copyright issue, our dataset consists of three parts:

- 1403 editions in 670 languages with permissive licenses which we distribute freely (the corpus we call Taxi1500-c v1.0).
- For the remaining PBC Bibles, please contact Michael Cysouw at Philipps University of Marburg to request access to PBC. Once granted access, run the code available at our Github to obtain the labeled dataset.
- For the remaining 1000Langs Bibles, use the code provided at the corresponding Github to crawl the corpus. Then, run the code available at our Github to obtain the labeled dataset.

G Details of Taxi1500 dataset

We represent the definition of every class in Table 8. and the number of verses of different languages in table 9.

H Evaluation on LLMs

We use a 3-shot prompt and adhere to the methodology outlined in Lin et al. (2024). We report average results in 5 and all results in 10.

I Results for zero-shot

We report the detailed results for zero-shot transfer of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

class	definition
Recommendation	`An imperative statement which suggests to act or believe in certain ways.`
Faith	`Display of belief and love toward God, instructions on how to maintain faith, stories of faith and its consequences, etc.`
Description	`Describes a person, relationship, phenomenon, situation, etc.`
Sin	`Describes what is considered sin, stories of sinful people and sinful actions.`
Grace	`God's love, blessing, and kindness towards humans.`
Violence	`Describes wars, conflict, threats, and torture; but also destructions of people, cities, and nations.`

Table 8: Definitions of the six Taxi1500 classes

verse.num	1077	1076	1075	1074	1073	1072	1071	1070	1069	1067	1066	1065
lan.num	1409	20	14	5	4	2	3	5	1	2	2	3
verse.num	1064	1063	1061	1060	1057	1056	1055	1054	1053	1051	1049	1048
lan.num	3	1	2	3	1	2	3	1	1	1	1	3
verse.num	1044	1042	1041	1039	1038	1034	1017	1006	1000	989	961	949
lan.num	1	1	1	1	1	1	1	2	1	1	1	1

Table 9: An overview of the number of verses of different languages, for example: 1049 of the languages have 1077 verses in the dataset.

Language	LLaMA2	Mistral	Bloom-560M	Bloom-1B	Bloom-3B	Bloom-7B
mhr_Cyrl	0.47	0.46	0.48	0.50	**0.51**	0.46
azb_Arab	0.40	**0.51**	0.43	0.47	0.41	0.48
asm_Beng	0.46	**0.56**	0.36	0.45	0.49	0.55
ben_Beng	0.41	**0.58**	0.41	0.48	0.48	0.52
tha_Thai	0.43	**0.58**	0.45	0.47	0.41	0.43
khm_Khmr	0.52	**0.56**	0.52	**0.56**	0.52	0.49
ell_Grek	0.49	**0.58**	0.44	0.49	0.49	0.49
oss_Cyrl	0.49	0.48	0.48	**0.52**	0.47	0.49
pan_Guru	0.41	0.46	0.44	**0.47**	**0.47**	**0.47**
tat_Cyrl	0.48	**0.53**	0.43	**0.53**	0.48	0.46
hne_Deva	0.56	**0.61**	0.56	**0.61**	0.58	0.54
arb_Arab	0.43	**0.62**	0.46	0.53	0.49	0.49
mkd_Cyrl	0.52	**0.67**	0.54	0.61	0.57	0.57
bul_Cyrl	0.45	**0.61**	0.41	0.44	0.47	0.49
kir_Cyrl	0.51	0.53	**0.62**	**0.62**	0.57	0.48
kaz_Cyrl	0.49	**0.55**	0.45	0.51	**0.55**	0.51
udm_Cyrl	0.37	0.41	0.42	**0.45**	0.43	0.42
kat_Geor	0.41	**0.45**	0.43	**0.45**	0.43	0.42
sah_Cyrl	0.41	0.46	**0.49**	**0.49**	0.46	0.44
mai_Deva	0.45	**0.62**	0.45	0.52	0.49	0.49
ary_Arab	0.32	**0.56**	0.34	0.43	0.36	0.39
tyv_Cyrl	0.39	**0.48**	0.36	0.45	**0.48**	0.43
snd_Arab	0.44	**0.62**	0.54	0.56	0.49	0.57
tir_Ethi	0.30	0.40	0.38	**0.41**	0.32	0.28
mya_Mymr	0.45	0.51	0.51	**0.53**	0.41	0.44
alt_Cyrl	0.44	0.46	0.49	**0.53**	0.48	0.45
fas_Arab	0.49	**0.67**	0.53	0.53	0.49	0.58
kor_Hang	0.49	**0.72**	0.49	0.51	0.52	0.49
krc_Cyrl	0.46	**0.55**	0.45	0.49	0.46	0.49
mar_Deva	0.49	**0.56**	0.49	0.49	0.49	0.53
chv_Cyrl	0.43	0.45	0.47	**0.51**	0.42	0.45
crh_Cyrl	0.49	**0.57**	0.48	0.49	0.51	0.48
npi_Deva	0.51	**0.67**	0.56	0.55	0.59	0.56
pes_Arab	0.51	**0.65**	0.54	0.50	0.49	0.59
nep_Deva	0.45	**0.67**	0.51	0.58	0.54	0.63
hin_Deva	0.51	**0.65**	0.55	0.48	0.47	0.49
arz_Arab	0.32	**0.54**	0.35	0.44	0.41	0.45
ksw_Mymr	0.44	0.44	0.40	**0.49**	0.42	0.42
rus_Cyrl	0.49	**0.58**	0.43	0.47	0.45	0.51
bel_Cyrl	0.48	**0.56**	0.46	0.51	0.45	0.49
ckb_Arab	0.44	**0.48**	0.45	0.47	0.43	0.45
lao_Laoo	0.45	0.45	0.48	0.51	**0.57**	0.47
tgk_Cyrl	0.42	**0.56**	0.46	0.54	0.48	0.49
lzh_Hani	0.55	**0.66**	0.51	0.56	0.53	0.54
tel_Telu	0.33	**0.54**	0.39	0.52	0.51	0.51
sin_Sinh	0.40	0.38	0.41	**0.47**	0.42	0.40
prs_Arab	0.51	**0.66**	0.57	0.60	0.57	0.56
che_Cyrl	0.38	**0.42**	0.36	0.41	0.33	0.37
uzn_Cyrl	0.46	**0.59**	0.43	0.49	0.43	0.45
myv_Cyrl	0.40	0.45	0.36	**0.47**	0.45	0.41
tam_Taml	0.44	**0.60**	0.55	0.55	**0.60**	0.59
cmn_Hani	0.49	**0.61**	0.44	0.54	0.54	0.53
kjh_Cyrl	0.44	0.48	0.42	**0.49**	0.42	0.45
hye_Armn	0.46	**0.55**	0.46	0.52	0.52	0.46
bak_Cyrl	0.45	0.49	0.45	**0.51**	0.47	0.49
kmr_Cyrl	0.40	0.40	0.39	0.44	0.43	**0.45**
mdy_Ethi	0.40	**0.55**	0.47	0.46	0.45	0.43
ukr_Cyrl	0.52	**0.63**	0.51	0.49	0.49	0.51
suz_Deva	0.47	0.43	0.42	**0.48**	0.45	0.42
guj_Gujr	0.46	**0.52**	0.46	0.48	0.51	**0.52**
dzo_Tibt	**0.45**	**0.45**	0.42	0.41	0.43	0.41
ori_Orya	0.43	0.51	0.51	**0.56**	0.54	0.51
ory_Orya	0.44	0.58	0.53	0.51	**0.59**	0.49
yue_Hani	0.43	**0.63**	0.46	0.54	0.53	0.53

Table 10: Performance across six LLMs on 64 selected languages.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
aah_Latn	**0.13**	0.10	0.05	0.05	0.08
aai_Latn	**0.22**	0.15	0.09	0.05	0.09
aak_Latn	0.07	**0.13**	0.05	0.05	0.05
aau_Latn	0.12	**0.12**	0.06	0.05	0.10
aaz_Latn	0.07	**0.12**	0.05	0.05	0.08
abi_Latn	0.07	**0.11**	0.05	0.05	0.05
abt_Latn	0.09	**0.13**	0.08	0.05	0.06
abx_Latn	0.16	0.12	0.20	0.14	**0.33**
aby_Latn	**0.21**	0.12	0.07	0.07	0.06
acd_Latn	**0.13**	0.08	0.05	0.05	0.05
ace_Latn	0.13	0.25	0.11	0.11	**0.30**
acf_Latn	0.09	0.25	0.06	0.05	**0.38**
ach_Latn	**0.13**	0.12	0.05	0.05	0.08
acn_Latn	0.07	**0.10**	0.05	0.05	0.05
acr_Latn	0.16	0.14	0.06	0.05	**0.30**
acu_Latn	0.10	**0.10**	0.05	0.05	0.08
ade_Latn	**0.12**	0.10	0.07	0.05	0.06
adh_Latn	0.13	**0.15**	0.07	0.05	0.07
adi_Latn	0.09	0.10	**0.14**	0.05	0.09
adj_Latn	**0.17**	0.08	0.05	0.05	0.05
adl_Latn	0.08	**0.18**	0.05	0.05	0.05
aeb_Arab	0.07	0.38	0.19	**0.42**	0.30
aer_Latn	0.07	**0.08**	0.08	0.05	0.05
aeu_Latn	0.07	**0.13**	0.05	0.05	0.05
aey_Latn	0.07	**0.12**	0.09	0.05	0.05
afr_Latn	0.33	0.45	0.59	**0.66**	0.52
agd_Latn	0.09	**0.16**	0.06	0.08	0.07
agg_Latn	**0.14**	0.06	0.05	0.05	0.05
agm_Latn	0.07	**0.11**	0.06	0.05	0.05
agn_Latn	0.12	0.16	0.13	0.18	**0.35**
agr_Latn	0.07	**0.11**	0.05	0.05	0.05
agt_Latn	0.07	**0.10**	0.06	0.05	0.10
agu_Latn	**0.11**	0.09	0.04	0.05	0.06
agw_Latn	0.20	0.13	0.11	0.07	**0.24**
ahk_Latn	0.08	**0.11**	0.07	0.05	0.07
aia_Latn	**0.23**	0.13	0.05	0.05	0.08
aii_Syrc	0.07	0.05	0.05	0.09	**0.10**
aim_Latn	0.10	**0.14**	0.06	0.05	0.05
ain_Latn	**0.11**	0.09	0.07	0.05	0.10
aji_Latn	0.13	**0.14**	0.05	0.05	0.05
ajz_Latn	0.12	**0.12**	0.05	0.05	0.07
aka_Latn	0.12	**0.17**	0.10	0.06	0.13
akb_Latn	0.13	0.16	0.15	0.07	**0.27**
ake_Latn	**0.11**	0.08	0.05	0.05	0.05
akh_Latn	0.10	**0.15**	0.05	0.05	0.05
akp_Latn	0.10	**0.16**	0.06	0.05	0.05
ald_Latn	**0.08**	0.05	0.05	0.05	0.05
alj_Latn	0.11	0.14	0.10	0.10	**0.21**
aln_Latn	0.07	0.25	0.46	0.53	**0.55**
alp_Latn	0.10	0.19	0.13	0.06	**0.20**
alq_Latn	0.09	**0.11**	0.05	0.05	0.05
als_Latn	0.07	0.24	0.45	**0.54**	0.49
alt_Cyrl	0.07	0.16	0.17	0.19	**0.37**
alz_Latn	0.10	0.15	0.06	0.05	**0.17**
ame_Latn	0.09	**0.11**	0.09	0.05	0.05
amf_Latn	0.07	**0.08**	0.05	0.05	0.05
amh_Ethi	0.07	0.05	**0.10**	0.05	0.07
amk_Latn	0.13	**0.19**	0.06	0.05	0.07
amm_Latn	**0.09**	0.07	0.04	0.05	0.08
amn_Latn	0.11	0.11	0.07	0.05	**0.12**
amp_Latn	0.07	**0.12**	0.06	0.05	0.05
amr_Latn	0.09	**0.12**	0.05	0.05	0.05
amu_Latn	0.06	**0.08**	0.05	0.05	0.05
anm_Latn	0.13	**0.14**	0.06	0.05	0.05
ann_Latn	0.14	**0.15**	0.08	0.05	0.06
anv_Latn	0.13	**0.13**	0.05	0.05	0.08
any_Latn	0.07	**0.07**	0.05	0.05	0.05
aoj_Latn	**0.20**	0.09	0.08	0.05	0.06
aom_Latn	**0.23**	0.16	0.05	0.05	0.05
aon_Latn	0.08	**0.11**	0.06	0.05	0.05
aoz_Latn	**0.21**	0.13	0.07	0.05	0.07
apb_Latn	0.07	0.08	0.06	0.05	**0.12**
ape_Latn	0.13	**0.13**	0.05	0.05	0.07
apn_Latn	0.07	**0.19**	0.06	0.05	0.05
apr_Latn	0.07	**0.07**	0.07	0.05	0.05
apt_Latn	0.08	**0.14**	0.07	0.05	0.07
apu_Latn	0.07	0.09	**0.10**	0.05	0.05
apw_Latn	**0.15**	0.10	0.05	0.05	0.05
apy_Latn	0.09	0.09	**0.11**	0.05	0.05
apz_Latn	0.07	**0.11**	0.05	0.05	0.05
are_Latn	0.11	**0.12**	0.05	0.05	0.05
arl_Latn	**0.15**	0.14	0.05	0.05	0.05
arn_Latn	**0.13**	0.08	0.05	0.05	0.08
ary_Arab	0.07	**0.28**	0.19	0.27	0.19
arz_Arab	0.07	0.43	0.32	**0.47**	0.25
asg_Latn	0.08	**0.11**	0.05	0.05	0.06
asm_Beng	0.07	0.17	0.43	0.47	**0.51**
aso_Latn	**0.15**	0.12	0.05	0.05	0.05
ata_Latn	0.11	**0.12**	0.06	0.05	0.06
atb_Latn	**0.10**	0.09	0.07	0.05	0.06
atd_Latn	**0.11**	0.09	0.05	0.05	0.05
atg_Latn	0.10	**0.11**	0.07	0.05	0.07
atq_Latn	0.13	**0.15**	0.06	0.05	0.13
att_Latn	0.14	0.10	0.08	0.05	**0.16**
auc_Latn	0.09	**0.13**	0.06	0.05	0.05
auy_Latn	0.07	**0.07**	0.04	0.05	0.06
ava_Cyrl	0.07	0.06	0.05	0.05	**0.10**
avn_Latn	**0.14**	0.12	0.05	0.05	0.05
avt_Latn	0.10	0.11	0.05	0.05	**0.14**
avu_Latn	**0.07**	0.06	0.04	0.05	0.05
awa_Deva	0.07	0.24	0.37	0.40	**0.48**
awb_Latn	0.08	**0.11**	0.06	0.05	0.05
awi_Latn	**0.17**	0.12	0.04	0.05	0.14
ayo_Latn	0.12	**0.12**	0.10	0.05	0.08
ayp_Arab	0.07	0.30	0.29	0.35	**0.43**
ayr_Latn	0.07	**0.12**	0.11	0.06	0.10
azb_Arab	0.07	0.16	0.15	0.08	**0.34**
aze_Latn	0.07	0.32	0.56	**0.68**	0.59
azg_Latn	0.04	**0.09**	0.05	0.05	0.05
azz_Latn	0.14	**0.15**	0.06	0.06	0.10
bak_Cyrl	0.07	**0.33**	0.13	0.05	0.24
bam_Latn	0.09	0.11	0.06	0.05	**0.20**
ban_Latn	0.07	0.16	0.16	0.09	**0.31**
bao_Latn	0.10	**0.14**	0.08	0.05	0.06
bar_Latn	0.13	0.19	0.30	0.29	**0.41**
bav_Latn	**0.12**	0.05	0.05	0.05	0.06
bba_Latn	**0.13**	0.12	0.05	0.05	0.05
bbb_Latn	0.07	**0.09**	0.05	0.05	0.05
bbj_Latn	**0.12**	0.05	0.05	0.05	0.05
bbk_Latn	**0.09**	0.04	0.05	0.05	0.05
bbo_Latn	0.10	**0.12**	0.07	0.05	0.06
bbr_Latn	**0.17**	0.15	0.04	0.05	0.06
bch_Latn	0.10	**0.13**	0.07	0.05	0.12
bci_Latn	0.09	0.12	0.04	0.05	**0.15**
bcl_Latn	0.07	0.18	0.26	0.20	**0.46**
bcw_Latn	**0.12**	0.05	0.06	0.05	0.05
bdd_Latn	**0.11**	0.07	0.05	0.05	0.05
bdh_Latn	0.07	**0.10**	0.05	0.05	0.05
bdq_Latn	0.10	**0.12**	0.05	0.05	0.05
bef_Latn	0.10	**0.10**	0.07	0.05	0.07
bel_Cyrl	0.07	0.43	0.59	**0.67**	0.59
bem_Latn	0.14	0.11	0.08	0.09	**0.31**
ben_Beng	0.07	0.32	0.56	**0.67**	0.63
beq_Latn	0.14	**0.14**	0.09	0.05	0.10
bex_Latn	**0.13**	0.10	0.05	0.05	0.08
bfd_Latn	**0.11**	0.09	0.05	0.05	0.05
bfo_Latn	0.10	**0.11**	0.05	0.05	0.06
bgr_Latn	0.16	0.17	0.07	0.05	**0.30**
bgs_Latn	**0.15**	0.14	0.09	0.07	0.11
bgt_Latn	0.15	**0.16**	0.07	0.05	0.16

Table 11: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
bgz_Latn	0.09	**0.18**	0.09	0.06	0.15	bzj_Latn	0.24	0.15	0.13	0.06	**0.35**
bhl_Latn	0.10	**0.12**	0.06	0.05	0.07	caa_Latn	0.14	**0.15**	0.07	0.05	0.12
bhp_Latn	0.09	0.11	**0.16**	0.06	0.09	cab_Latn	0.07	**0.10**	0.05	0.05	0.05
bhw_Latn	0.09	**0.16**	0.07	0.05	0.14	cac_Latn	0.12	0.12	0.06	0.05	**0.21**
bhz_Latn	**0.18**	0.14	0.06	0.05	0.06	caf_Latn	**0.09**	0.07	0.05	0.05	0.05
bib_Latn	**0.16**	0.06	0.05	0.05	0.06	cag_Latn	0.07	**0.14**	0.05	0.05	0.11
big_Latn	0.09	**0.10**	0.05	0.05	0.05	cak_Latn	0.04	0.12	0.05	0.05	**0.42**
bim_Latn	**0.14**	0.13	0.05	0.05	0.06	cao_Latn	0.08	**0.10**	0.05	0.05	0.10
bis_Latn	0.16	0.22	0.14	0.06	**0.24**	cap_Latn	**0.11**	0.09	0.05	0.05	0.05
biu_Latn	0.16	0.14	0.05	0.05	**0.17**	caq_Latn	0.10	**0.10**	0.04	0.05	0.10
biv_Latn	**0.11**	0.07	0.05	0.05	0.05	car_Latn	**0.13**	0.12	0.06	0.05	0.06
bjr_Latn	0.07	**0.10**	0.05	0.05	0.05	cas_Latn	**0.15**	0.09	0.08	0.05	0.04
bjv_Latn	**0.11**	0.08	0.06	0.05	0.05	cat_Latn	0.13	0.41	0.58	**0.64**	0.47
bkd_Latn	0.07	**0.21**	0.15	0.08	0.21	cav_Latn	0.07	**0.11**	0.06	0.05	0.05
bkl_Latn	**0.15**	0.11	0.06	0.07	0.05	cax_Latn	0.07	**0.12**	0.09	0.05	0.06
bkq_Latn	**0.14**	0.12	0.06	0.05	0.11	cbc_Latn	0.08	**0.14**	0.06	0.05	0.05
bku_Latn	0.15	0.11	0.08	0.06	**0.19**	cbi_Latn	**0.14**	0.13	0.09	0.05	0.11
bkv_Latn	**0.13**	0.06	0.06	0.05	0.09	cbk_Latn	0.11	0.39	0.45	0.48	**0.57**
blh_Latn	0.05	**0.07**	0.05	0.05	0.05	cbr_Latn	0.13	**0.15**	0.05	0.05	0.05
blt_Latn	**0.11**	0.08	0.07	0.05	0.06	cbs_Latn	0.05	**0.15**	0.05	0.05	0.06
blw_Latn	0.07	**0.15**	0.06	0.05	0.10	cbt_Latn	0.08	**0.09**	0.06	0.05	0.06
blz_Latn	0.15	**0.19**	0.09	0.06	0.12	cbu_Latn	0.07	**0.12**	0.05	0.05	0.05
bmb_Latn	0.14	**0.14**	0.09	0.05	0.10	cbv_Latn	0.09	**0.15**	0.06	0.05	0.08
bmh_Latn	0.07	**0.11**	0.08	0.05	0.08	cce_Latn	0.09	0.10	0.09	0.05	**0.21**
bmq_Latn	**0.10**	0.07	0.05	0.05	0.05	cco_Latn	**0.10**	0.06	0.05	0.05	0.05
bmr_Latn	0.07	**0.13**	0.05	0.05	0.05	ccp_Latn	0.11	**0.19**	0.09	0.06	0.09
bmu_Latn	0.09	**0.14**	0.05	0.05	0.05	cdf_Latn	0.09	**0.12**	0.05	0.05	0.09
bmv_Latn	**0.16**	0.10	0.07	0.05	0.05	ceb_Latn	0.11	0.12	0.28	0.28	**0.37**
bnj_Latn	0.09	**0.13**	0.07	0.06	0.05	ceg_Latn	0.15	**0.15**	0.04	0.05	0.08
bno_Latn	0.10	0.18	0.18	0.11	**0.33**	cek_Latn	0.09	**0.10**	0.05	0.05	0.06
bnp_Latn	0.11	0.13	0.05	0.06	**0.16**	ces_Latn	0.07	0.28	**0.66**	0.57	0.51
boa_Latn	0.09	**0.16**	0.05	0.05	0.05	cfm_Latn	0.14	0.15	0.05	0.05	**0.25**
boj_Latn	**0.13**	0.10	0.05	0.05	0.07	cgc_Latn	0.07	0.18	0.19	0.14	**0.26**
bom_Latn	0.08	**0.11**	0.05	0.05	0.08	cha_Latn	0.12	0.12	0.11	0.05	**0.19**
bon_Latn	0.11	**0.19**	0.07	0.06	0.05	chd_Latn	0.09	**0.10**	0.05	0.05	0.06
bov_Latn	0.07	**0.12**	0.05	0.05	0.06	che_Cyrl	0.07	**0.10**	0.07	0.05	0.08
box_Latn	0.09	**0.11**	0.05	0.05	0.09	chf_Latn	0.09	0.10	0.12	0.05	**0.21**
bpr_Latn	0.13	**0.13**	0.09	0.05	0.09	chj_Latn	**0.10**	0.06	0.05	0.05	0.05
bps_Latn	**0.16**	0.11	0.08	0.05	0.08	chk_Hani	0.07	**0.13**	0.07	0.05	0.08
bqc_Latn	0.07	**0.11**	0.05	0.05	0.06	chq_Latn	0.09	**0.10**	0.05	0.05	0.05
bqj_Latn	**0.17**	0.12	0.09	0.05	0.07	chr_Cher	0.07	0.05	**0.09**	0.05	0.05
bqp_Latn	0.09	**0.17**	0.05	0.05	0.06	chu_Cyrl	0.07	0.31	0.60	**0.61**	0.46
bre_Latn	0.08	0.29	0.25	**0.43**	0.29	chv_Cyrl	0.07	0.18	0.07	0.05	**0.19**
bru_Latn	0.10	**0.10**	0.07	0.05	0.05	chz_Latn	0.07	**0.08**	0.05	0.05	0.05
bsc_Latn	**0.15**	0.08	0.09	0.05	0.05	cjo_Latn	0.07	**0.07**	0.04	0.05	0.05
bsn_Latn	**0.16**	0.07	0.04	0.05	0.07	cjp_Latn	**0.14**	0.11	0.07	0.05	0.05
bss_Latn	0.07	**0.13**	0.10	0.05	0.05	cjv_Latn	0.06	**0.08**	0.07	0.05	0.05
btd_Latn	0.09	**0.30**	0.21	0.17	0.28	ckb_Latn	0.16	0.09	0.07	0.07	**0.43**
bth_Latn	0.10	0.14	0.12	0.07	**0.25**	cko_Latn	0.08	**0.09**	0.06	0.05	0.06
bto_Latn	0.07	0.11	0.13	0.05	**0.32**	cle_Latn	**0.11**	0.04	0.05	0.05	0.06
btt_Latn	0.12	**0.14**	0.07	0.05	0.06	clu_Latn	0.11	0.14	0.18	0.21	**0.43**
btx_Latn	0.16	0.23	0.20	0.19	**0.34**	cly_Latn	**0.15**	0.12	0.11	0.05	0.06
bud_Latn	0.05	**0.12**	0.05	0.05	0.05	cme_Latn	0.09	**0.12**	0.05	0.05	0.05
bug_Latn	0.09	**0.19**	0.12	0.07	0.17	cmn_Hani	0.07	0.40	0.59	0.62	**0.65**
buk_Latn	0.07	**0.11**	0.05	0.05	0.08	cmo_Latn	**0.18**	0.17	0.13	0.05	0.05
bul_Cyrl	0.07	0.41	0.62	**0.64**	0.60	cmr_Latn	0.11	**0.13**	0.05	0.05	0.06
bum_Latn	0.09	0.16	0.06	0.05	**0.17**	cnh_Latn	0.18	0.12	0.08	0.05	**0.20**
bus_Latn	0.08	**0.13**	0.05	0.05	0.05	cni_Latn	0.07	**0.07**	0.05	0.05	0.05
bvc_Latn	0.14	**0.21**	0.06	0.05	0.08	cnk_Latn	0.09	**0.09**	0.05	0.05	0.06
bvd_Latn	**0.19**	0.11	0.06	0.05	0.08	cnl_Latn	0.07	**0.07**	0.05	0.05	0.05
bvr_Latn	**0.12**	0.07	0.09	0.05	0.05	cnt_Latn	0.07	**0.08**	0.05	0.05	0.05
bvz_Latn	**0.13**	0.10	0.08	0.05	0.05	cnw_Latn	0.12	0.13	0.06	0.05	**0.14**
bwq_Latn	**0.15**	0.09	0.06	0.05	0.11	coe_Latn	0.07	**0.08**	0.05	0.05	0.06
bwu_Latn	0.14	**0.16**	0.08	0.05	0.09	cof_Latn	0.11	**0.15**	0.06	0.05	0.08
bxr_Cyrl	0.07	0.09	0.25	0.27	**0.31**	cok_Latn	**0.13**	0.08	0.05	0.05	0.07
byr_Latn	0.07	**0.08**	0.05	0.05	0.06	con_Latn	**0.28**	0.07	0.10	0.05	0.07
byx_Latn	0.07	**0.13**	0.07	0.06	0.05	cop_Copt	0.07	**0.07**	0.05	0.05	0.05
bzd_Latn	0.07	**0.10**	0.05	0.05	0.04	cor_Latn	0.09	**0.12**	0.09	0.05	0.11
bzh_Latn	**0.15**	0.08	0.05	0.05	0.05	cot_Latn	0.07	**0.12**	0.05	0.05	0.05
bzi_Thai	0.07	**0.07**	0.07	0.05	0.05	cou_Latn	0.10	**0.14**	0.06	0.05	0.05

Table 12: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
cpa_Latn	0.07	**0.11**	0.05	0.05	0.05	due_Latn	0.10	0.12	0.16	0.05	**0.20**
cpb_Latn	0.07	**0.08**	0.08	0.05	0.05	dug_Latn	0.08	**0.17**	0.17	0.11	0.16
cpc_Latn	0.09	**0.12**	0.06	0.05	0.05	duo_Latn	0.14	0.08	0.16	0.06	**0.31**
cpu_Latn	0.09	**0.11**	0.04	0.07	0.05	dur_Latn	0.10	**0.10**	0.05	0.05	0.05
cpy_Latn	0.07	**0.08**	0.05	0.05	0.05	dwr_Latn	**0.15**	0.11	0.06	0.05	0.10
crh_Cyrl	0.07	0.19	0.15	0.20	**0.45**	dww_Latn	0.07	0.07	**0.08**	0.05	0.06
crj_Latn	**0.15**	0.10	0.05	0.05	0.05	dyi_Latn	**0.16**	0.13	0.07	0.05	0.06
crk_Cans	**0.07**	0.05	0.05	0.05	0.05	dyo_Latn	0.08	**0.12**	0.07	0.05	0.08
crl_Cans	0.07	**0.09**	0.05	0.05	0.05	dyu_Latn	0.07	0.09	0.05	0.05	**0.17**
crm_Cans	**0.07**	0.05	0.05	0.05	0.06	dzo_Tibt	0.07	0.04	0.05	0.08	**0.09**
crn_Latn	**0.10**	0.09	0.05	0.05	0.06	ebk_Latn	0.14	0.15	0.05	0.05	**0.17**
crq_Latn	0.09	**0.16**	0.06	0.05	0.05	efi_Latn	0.13	**0.13**	0.07	0.05	0.11
crs_Latn	0.10	0.17	0.15	0.05	**0.43**	eka_Latn	0.11	**0.17**	0.09	0.06	0.06
crt_Latn	0.10	**0.16**	0.06	0.05	0.05	ell_Grek	0.07	0.31	0.43	**0.60**	0.50
crx_Latn	**0.09**	0.08	0.08	0.05	0.05	emi_Latn	0.09	**0.16**	0.05	0.10	0.09
csk_Latn	0.12	**0.14**	0.09	0.05	0.05	emp_Latn	**0.14**	0.10	0.06	0.05	0.05
cso_Latn	0.07	**0.08**	0.05	0.05	0.05	enb_Latn	0.07	**0.10**	0.05	0.05	0.05
csy_Latn	0.10	0.11	0.08	0.05	**0.14**	eng_Latn	0.43	**0.71**	0.65	0.56	0.63
cta_Latn	0.07	**0.13**	0.05	0.05	0.07	enl_Latn	0.09	**0.10**	0.05	0.05	0.07
ctd_Latn	0.11	0.14	0.07	0.05	**0.22**	enm_Latn	0.33	0.46	0.55	0.45	**0.55**
ctp_Latn	**0.14**	0.08	0.06	0.05	0.06	enq_Latn	0.07	**0.12**	0.05	0.05	0.07
ctu_Latn	0.10	0.09	0.11	0.06	**0.27**	epo_Latn	0.15	0.25	0.57	**0.61**	0.48
cub_Latn	**0.11**	0.08	0.05	0.05	0.05	eri_Latn	0.13	**0.13**	0.07	0.06	0.06
cuc_Latn	0.07	**0.13**	0.05	0.05	0.05	ese_Latn	0.09	**0.13**	0.06	0.05	0.06
cui_Latn	0.08	**0.14**	0.05	0.05	0.05	esi_Latn	**0.21**	0.12	0.05	0.05	0.07
cuk_Latn	**0.16**	0.11	0.13	0.05	0.07	esk_Latn	0.07	**0.11**	0.05	0.05	0.05
cul_Latn	0.09	**0.12**	0.07	0.05	0.05	ess_Latn	**0.14**	0.13	0.06	0.05	0.05
cut_Latn	**0.11**	0.10	0.05	0.05	0.07	est_Latn	0.07	0.46	**0.68**	0.56	0.47
cux_Latn	**0.16**	0.14	0.05	0.06	0.08	esu_Latn	**0.16**	0.12	0.05	0.05	0.05
cwe_Latn	0.11	0.19	0.13	0.11	**0.22**	etu_Latn	**0.13**	0.11	0.05	0.05	0.05
cwt_Latn	0.09	**0.14**	0.05	0.05	0.05	eus_Latn	0.09	0.18	**0.26**	0.25	0.23
cya_Latn	0.12	0.11	**0.14**	0.05	0.11	ewe_Latn	0.11	**0.11**	0.05	0.05	0.07
cym_Latn	0.08	0.23	0.44	**0.53**	0.49	ewo_Latn	0.13	**0.18**	0.08	0.06	0.10
czt_Latn	**0.14**	0.11	0.07	0.05	0.05	eza_Latn	0.07	**0.09**	0.05	0.05	0.06
daa_Latn	**0.13**	0.09	0.06	0.06	0.05	faa_Latn	**0.11**	0.08	0.07	0.05	0.08
dad_Latn	**0.20**	0.15	0.06	0.05	0.05	fai_Latn	**0.13**	0.11	0.06	0.05	0.05
dah_Latn	0.12	**0.17**	0.05	0.05	0.05	fal_Latn	**0.20**	0.15	0.09	0.05	0.06
dan_Latn	0.19	0.52	**0.54**	0.54	0.53	fao_Latn	0.09	0.27	0.32	0.36	**0.48**
dbq_Latn	**0.13**	0.07	0.06	0.05	0.05	far_Latn	0.20	**0.20**	0.07	0.06	0.14
ddn_Latn	**0.10**	0.05	0.10	0.05	0.05	fas_Arab	0.07	0.46	0.67	0.66	**0.67**
ded_Latn	0.07	**0.09**	0.06	0.05	0.06	ffm_Latn	**0.13**	0.11	0.05	0.05	0.07
des_Latn	0.07	**0.10**	0.05	0.05	0.05	fij_Latn	0.05	**0.12**	0.08	0.05	0.12
deu_Latn	0.15	0.38	**0.52**	0.52	0.46	fil_Latn	0.13	0.29	0.47	0.55	**0.55**
dga_Latn	0.10	**0.13**	0.05	0.05	0.05	fin_Latn	0.13	0.45	**0.58**	0.57	0.47
dgc_Latn	0.16	0.14	0.21	0.18	**0.25**	fon_Latn	**0.10**	0.09	0.05	0.05	0.05
dgi_Latn	**0.12**	0.07	0.05	0.05	0.06	for_Latn	0.09	**0.12**	0.07	0.05	0.06
dgr_Latn	0.10	**0.11**	0.05	0.05	0.05	fra_Latn	0.13	0.54	**0.65**	0.65	0.54
dgz_Latn	**0.20**	0.13	0.12	0.06	0.15	frd_Latn	0.08	**0.13**	0.06	0.05	0.09
dhm_Latn	0.17	**0.17**	0.10	0.05	0.10	fry_Latn	0.21	0.38	0.30	0.37	**0.42**
did_Latn	0.07	**0.14**	0.05	0.05	0.05	fub_Latn	**0.17**	0.16	0.10	0.05	0.12
dig_Latn	0.12	0.14	0.20	0.23	**0.39**	fue_Latn	0.13	**0.14**	0.07	0.05	0.14
dik_Latn	**0.12**	0.09	0.08	0.05	0.06	fuf_Latn	0.10	0.10	0.09	0.05	**0.13**
dip_Latn	0.15	**0.15**	0.05	0.05	0.06	fuh_Latn	**0.12**	0.09	0.05	0.06	0.05
dis_Latn	**0.13**	0.11	0.10	0.05	0.06	fuq_Latn	0.11	**0.11**	0.10	0.05	0.10
dje_Latn	**0.12**	0.09	0.08	0.05	0.07	fuv_Latn	0.11	0.13	0.11	0.05	**0.14**
djk_Latn	0.14	0.14	0.08	0.05	**0.28**	gaa_Latn	0.12	**0.13**	0.05	0.05	0.05
djr_Latn	0.07	**0.12**	0.05	0.05	0.05	gag_Latn	0.07	0.13	0.33	0.38	**0.40**
dks_Latn	**0.14**	0.12	0.05	0.05	0.05	gah_Latn	0.07	**0.15**	0.05	0.05	0.05
dln_Latn	0.12	0.12	0.05	0.05	**0.29**	gai_Latn	0.07	**0.09**	0.05	0.05	0.05
dnj_Latn	**0.10**	0.06	0.05	0.05	0.05	gam_Latn	**0.20**	0.11	0.11	0.05	0.11
dnw_Latn	**0.18**	0.12	0.07	0.05	0.06	gaw_Latn	**0.11**	0.09	0.06	0.05	0.08
dob_Latn	0.08	0.08	**0.10**	0.05	0.07	gbi_Latn	0.10	**0.11**	0.06	0.05	0.08
dop_Latn	**0.12**	0.07	0.05	0.05	0.05	gbo_Latn	0.08	**0.14**	0.05	0.05	0.05
dos_Latn	0.13	**0.14**	0.05	0.05	0.05	gbr_Latn	**0.17**	0.08	0.10	0.05	0.09
dow_Latn	0.06	**0.07**	0.05	0.05	0.05	gde_Latn	**0.10**	0.05	0.06	0.05	0.05
dru_Latn	0.07	**0.14**	0.09	0.05	0.09	gdg_Latn	0.10	**0.18**	0.09	0.06	0.16
dsh_Latn	**0.12**	0.10	0.07	0.05	0.06	gdn_Latn	0.07	**0.16**	0.07	0.06	0.09
dtb_Latn	0.11	**0.13**	0.06	0.05	0.08	gdr_Latn	**0.17**	0.09	0.05	0.05	0.06
dtp_Latn	0.12	0.12	0.05	0.05	**0.24**	geb_Latn	0.07	**0.08**	0.05	0.05	0.05
dts_Latn	0.09	**0.09**	0.05	0.05	0.06	gej_Latn	0.09	**0.10**	0.05	0.05	0.08

Table 13: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
gfk_Latn	**0.17**	0.12	0.07	0.05	0.10	hlt_Latn	0.09	**0.09**	0.05	0.05	0.06
ghe_Deva	0.07	0.11	0.20	0.15	**0.28**	hmo_Latn	0.09	**0.14**	0.09	0.05	0.07
ghs_Latn	0.07	**0.10**	0.05	0.05	0.06	hmr_Latn	**0.21**	0.06	0.07	0.05	0.20
gid_Latn	**0.10**	0.05	0.05	0.05	0.08	hne_Deva	0.07	0.27	0.29	0.39	**0.60**
gil_Latn	0.07	0.08	0.04	0.05	**0.23**	hnj_Latn	0.06	**0.06**	0.06	0.05	0.05
giz_Latn	0.07	**0.14**	0.06	0.05	0.07	hnn_Latn	0.11	0.17	0.17	0.12	**0.31**
gjn_Latn	0.09	**0.13**	0.05	0.05	0.05	hns_Latn	0.13	0.12	0.14	0.12	**0.19**
gkn_Latn	0.09	**0.16**	0.05	0.05	0.14	hop_Latn	**0.19**	0.17	0.05	0.05	0.11
gkp_Latn	0.09	**0.12**	0.05	0.05	0.07	hot_Latn	**0.11**	0.10	0.05	0.05	0.06
gla_Latn	0.12	0.14	0.34	0.42	**0.48**	hra_Latn	0.13	0.13	0.07	0.05	**0.26**
gle_Latn	0.17	0.15	0.38	**0.56**	0.40	hrv_Latn	0.09	0.35	0.64	**0.66**	0.63
glv_Latn	0.11	0.10	0.09	0.05	**0.11**	hto_Latn	**0.07**	0.06	0.05	0.06	0.05
gmv_Latn	**0.15**	0.12	0.07	0.06	0.06	hub_Latn	0.07	**0.13**	0.06	0.05	0.06
gna_Latn	0.11	**0.13**	0.05	0.05	0.05	hui_Latn	0.06	**0.10**	0.07	0.05	0.06
gnb_Latn	0.13	0.11	0.06	0.05	**0.20**	hun_Latn	0.08	0.38	**0.70**	0.66	0.52
gnd_Latn	**0.09**	0.06	0.05	0.05	0.05	hus_Latn	0.18	0.17	0.10	0.06	**0.20**
gng_Latn	0.12	**0.13**	0.06	0.05	0.05	huu_Latn	0.07	**0.11**	0.06	0.05	0.06
gnn_Latn	0.07	**0.10**	0.05	0.05	0.08	huv_Latn	0.07	**0.13**	0.06	0.05	0.11
gnw_Latn	0.07	**0.11**	0.07	0.05	0.06	hvn_Latn	0.14	**0.17**	0.09	0.05	0.11
gof_Latn	**0.15**	0.09	0.06	0.05	0.09	hwc_Latn	0.32	0.32	0.40	**0.53**	0.42
gog_Latn	0.13	0.13	0.11	0.07	**0.19**	hye_Armn	0.07	0.39	0.60	0.64	**0.65**
gom_Latn	0.07	0.11	0.06	0.05	**0.19**	ian_Latn	0.07	**0.12**	0.05	0.05	0.09
gor_Latn	0.12	0.17	0.08	0.09	**0.25**	iba_Latn	0.11	0.27	0.26	0.24	**0.54**
gqr_Latn	**0.19**	0.08	0.05	0.05	0.05	ibo_Latn	0.08	**0.12**	0.08	0.05	0.09
grt_Beng	0.07	0.10	**0.16**	0.05	0.11	icr_Latn	0.24	0.21	0.23	0.06	**0.40**
gso_Latn	0.07	**0.09**	0.05	0.05	0.05	ifa_Latn	0.10	0.15	0.06	0.05	**0.32**
gub_Latn	**0.13**	0.11	0.08	0.05	0.05	ifb_Latn	0.16	0.09	0.07	0.05	**0.32**
guc_Latn	0.13	**0.14**	0.05	0.05	0.05	ife_Latn	0.08	**0.11**	0.05	0.05	0.05
gud_Latn	0.11	**0.11**	0.05	0.05	0.05	ifk_Latn	0.14	0.14	0.07	0.05	**0.21**
gug_Latn	0.12	**0.17**	0.09	0.05	0.10	ifu_Latn	0.08	**0.17**	0.05	0.05	0.08
guh_Latn	0.07	**0.08**	0.06	0.05	0.06	ify_Latn	0.09	**0.14**	0.08	0.05	0.11
gui_Latn	0.09	**0.09**	0.09	0.05	0.07	ign_Latn	0.07	**0.09**	0.05	0.05	0.07
guj_Gujr	0.07	0.34	0.56	**0.70**	0.69	ike_Cans	0.07	0.05	0.05	0.05	**0.08**
guk_Ethi	0.07	0.10	0.07	0.05	**0.13**	ikk_Latn	0.07	**0.11**	0.11	0.05	0.05
gul_Latn	0.32	0.26	0.26	0.24	**0.49**	ikw_Latn	0.07	**0.07**	0.06	0.05	0.05
gum_Latn	0.07	**0.09**	0.05	0.05	0.06	ilb_Latn	0.09	0.12	0.14	0.09	**0.16**
gun_Latn	**0.12**	0.11	0.11	0.05	0.06	ilo_Latn	0.14	0.11	0.10	0.05	**0.33**
guo_Latn	0.13	0.09	0.08	0.06	**0.15**	imo_Latn	**0.14**	0.13	0.05	0.05	0.05
guq_Latn	0.07	0.15	**0.16**	0.05	0.06	inb_Latn	**0.11**	0.08	0.06	0.05	0.06
gur_Latn	0.13	**0.15**	0.05	0.05	0.09	ind_Latn	0.07	0.47	0.66	**0.70**	0.63
guu_Latn	**0.11**	0.10	0.06	0.05	0.06	ino_Latn	**0.14**	0.13	0.05	0.05	0.06
guw_Latn	**0.15**	0.12	0.11	0.05	0.05	iou_Latn	**0.14**	0.12	0.05	0.05	0.06
gux_Latn	0.07	**0.10**	0.07	0.05	0.07	ipi_Latn	0.07	**0.14**	0.04	0.05	0.05
guz_Latn	0.07	**0.15**	0.08	0.05	0.06	iqw_Latn	0.07	**0.12**	0.08	0.05	0.06
gvc_Latn	**0.14**	0.08	0.05	0.05	0.06	iri_Latn	0.12	**0.14**	0.05	0.05	0.05
gvf_Latn	**0.18**	0.09	0.06	0.05	0.06	irk_Latn	0.14	**0.15**	0.04	0.05	0.06
gvl_Latn	0.11	**0.14**	0.04	0.05	0.07	iry_Latn	0.08	0.14	0.11	0.16	**0.20**
gvn_Latn	0.07	**0.12**	0.05	0.05	0.09	isd_Latn	0.13	0.15	0.12	0.06	**0.19**
gwi_Latn	**0.19**	0.11	0.05	0.05	0.05	isl_Latn	0.07	0.33	0.57	**0.59**	0.47
gwr_Latn	**0.11**	0.10	0.08	0.05	0.09	ita_Latn	0.14	0.46	0.67	**0.68**	0.55
gya_Latn	0.10	**0.10**	0.05	0.05	0.06	itv_Latn	0.14	0.14	0.15	0.07	**0.27**
gym_Latn	0.11	0.09	**0.12**	0.05	0.07	ium_Latn	**0.10**	0.08	0.05	0.05	0.05
gyr_Latn	0.08	**0.10**	0.07	0.05	0.05	ivb_Latn	0.08	0.12	0.07	0.07	**0.17**
hae_Latn	0.09	0.15	0.15	**0.31**	0.22	ivv_Latn	0.11	0.13	0.07	0.05	**0.19**
hag_Latn	0.10	**0.13**	0.06	0.05	0.06	iws_Latn	**0.10**	0.09	0.05	0.05	0.05
hak_Latn	**0.13**	0.08	0.07	0.05	0.05	ixl_Latn	0.12	0.08	0.06	0.06	**0.16**
hat_Latn	0.06	0.17	0.08	0.06	**0.39**	izr_Latn	0.08	**0.14**	0.05	0.05	0.08
hau_Latn	0.14	0.15	0.36	**0.49**	0.40	izz_Latn	0.07	**0.13**	0.07	0.05	0.05
haw_Latn	0.12	0.11	0.05	0.05	**0.19**	jaa_Latn	0.10	**0.12**	0.06	0.05	0.08
hay_Latn	0.09	0.14	0.06	0.05	**0.15**	jac_Latn	**0.13**	0.07	0.06	0.05	0.09
hch_Latn	0.08	**0.13**	0.06	0.05	0.08	jae_Latn	0.07	**0.07**	0.05	0.05	0.05
heb_Hebr	0.07	**0.36**	0.15	0.31	0.24	jam_Latn	0.22	0.15	0.10	0.06	**0.46**
heg_Latn	0.07	**0.16**	0.05	0.05	0.09	jav_Latn	0.07	0.25	0.38	**0.57**	0.46
heh_Latn	0.10	**0.15**	0.11	0.09	0.09	jbu_Latn	0.12	**0.12**	0.08	0.05	0.08
hif_Latn	0.09	0.12	0.16	0.35	**0.43**	jic_Latn	0.13	**0.24**	0.07	0.05	0.12
hig_Latn	**0.15**	0.07	0.09	0.05	0.05	jiv_Latn	0.09	**0.15**	0.04	0.05	0.05
hil_Latn	0.14	0.23	0.26	0.24	**0.53**	jmc_Latn	**0.15**	0.10	0.05	0.06	0.09
hin_Deva	0.07	0.40	0.56	**0.62**	0.61	jpn_Jpan	0.07	0.37	**0.62**	0.56	0.50
hix_Latn	0.07	**0.08**	0.06	0.05	0.05	jra_Latn	0.09	**0.12**	0.06	0.05	0.06
hla_Latn	0.14	**0.15**	0.06	0.05	0.07	jun_Orya	0.07	0.05	0.11	0.06	**0.12**

Table 14: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
jvn_Latn	0.07	0.35	0.36	**0.52**	0.49	knf_Latn	0.13	**0.15**	0.07	0.05	0.05
kaa_Cyrl	0.07	0.17	0.14	0.16	**0.52**	kng_Latn	0.07	0.14	0.08	0.05	**0.15**
kab_Latn	0.11	**0.14**	0.07	0.06	0.13	knj_Latn	0.07	0.09	0.05	0.05	**0.18**
kac_Latn	**0.13**	0.10	0.05	0.05	0.05	knk_Latn	0.06	**0.11**	0.05	0.05	0.08
kal_Latn	0.09	0.11	0.05	0.05	**0.13**	kno_Latn	0.10	**0.10**	0.05	0.05	0.07
kan_Knda	0.07	0.34	0.56	**0.64**	0.61	knv_Latn	**0.18**	0.12	0.05	0.05	0.08
kao_Latn	0.09	**0.09**	0.05	0.05	0.06	kog_Latn	0.11	**0.12**	0.06	0.05	0.05
kaq_Latn	0.09	**0.16**	0.06	0.05	0.09	kor_Hang	0.07	0.43	0.63	**0.69**	0.62
kat_Geor	0.07	0.46	0.48	**0.61**	0.54	kpf_Latn	0.07	**0.10**	0.05	0.05	0.05
kaz_Cyrl	0.07	0.32	0.57	**0.66**	0.57	kpg_Latn	**0.22**	0.15	0.05	0.05	0.15
kbc_Latn	**0.18**	0.07	0.05	0.05	0.05	kpj_Latn	0.07	**0.10**	0.04	0.05	0.07
kbh_Latn	0.09	**0.13**	0.07	0.05	0.07	kpq_Latn	**0.15**	0.14	0.04	0.05	0.06
kbm_Latn	0.09	**0.15**	0.11	0.06	0.07	kpr_Latn	**0.13**	0.10	0.10	0.05	0.08
kbo_Latn	0.11	**0.15**	0.04	0.05	0.06	kpv_Cyrl	0.07	0.09	0.09	0.05	**0.11**
kbp_Latn	**0.10**	0.08	0.05	0.05	0.05	kpw_Latn	**0.14**	0.10	0.05	0.05	0.05
kbq_Latn	**0.12**	0.05	0.09	0.05	0.05	kpx_Latn	0.07	**0.13**	0.09	0.05	0.05
kbr_Latn	0.08	**0.13**	0.05	0.05	0.07	kpz_Latn	0.09	**0.12**	0.05	0.05	0.09
kcg_Latn	**0.13**	0.12	0.05	0.05	0.05	kqc_Latn	0.08	0.09	**0.11**	0.05	0.08
kck_Latn	0.08	0.13	0.09	0.05	**0.18**	kqe_Latn	0.13	0.16	0.13	0.12	**0.33**
kdc_Latn	0.13	0.14	0.20	0.19	**0.21**	kqo_Latn	0.07	**0.09**	0.05	0.05	0.05
kde_Latn	0.14	**0.16**	0.12	0.07	0.15	kqp_Latn	0.14	**0.14**	0.05	0.05	0.06
kdi_Latn	0.07	**0.16**	0.05	0.05	0.08	kqs_Latn	0.10	**0.13**	0.05	0.05	0.06
kdj_Latn	0.07	**0.13**	0.05	0.05	0.05	kqy_Ethi	0.07	**0.13**	0.06	0.05	0.05
kdl_Latn	0.07	**0.11**	0.07	0.05	0.09	krc_Cyrl	0.07	0.17	0.17	0.16	**0.48**
kdp_Latn	0.10	**0.11**	0.10	0.05	0.07	kri_Latn	0.15	0.16	0.05	0.05	**0.19**
kek_Latn	0.15	0.08	0.05	0.06	**0.27**	krj_Latn	0.11	0.21	0.33	0.28	**0.35**
ken_Latn	**0.10**	0.08	0.05	0.05	0.05	krl_Latn	0.07	0.34	0.40	0.40	**0.41**
keo_Latn	0.11	0.08	0.06	0.05	**0.11**	kru_Deva	0.07	**0.12**	0.08	0.05	0.11
ker_Latn	**0.09**	0.04	0.05	0.05	0.05	ksb_Latn	0.12	0.16	0.12	0.12	**0.21**
kew_Latn	0.13	**0.14**	0.05	0.05	0.06	ksc_Latn	0.09	**0.12**	0.07	0.05	0.11
kez_Latn	**0.13**	0.10	0.05	0.05	0.05	ksd_Latn	**0.15**	0.14	0.06	0.05	0.12
kff_Telu	0.07	0.14	**0.24**	0.20	0.20	ksf_Latn	**0.10**	0.07	0.05	0.05	0.06
kgf_Latn	0.08	**0.10**	0.05	0.05	0.05	ksr_Latn	0.08	**0.08**	0.05	0.05	0.06
kgk_Latn	0.07	**0.10**	0.06	0.05	0.05	kss_Latn	**0.12**	0.10	0.05	0.05	0.05
kgp_Latn	0.07	**0.14**	0.09	0.05	0.09	ksw_Mymr	0.07	**0.08**	0.05	0.05	0.06
kgr_Latn	0.14	**0.20**	0.06	0.05	0.13	ktb_Ethi	0.07	0.05	0.07	0.05	**0.10**
kha_Latn	**0.12**	0.07	0.07	0.05	0.06	ktj_Latn	0.04	**0.05**	0.05	0.05	0.05
khk_Latn	0.09	**0.15**	0.07	0.05	0.08	kto_Latn	0.07	**0.14**	0.09	0.05	0.05
khm_Khmr	0.07	0.05	0.55	**0.62**	0.55	ktu_Latn	0.10	0.11	0.11	0.06	**0.19**
khq_Latn	**0.12**	0.11	0.10	0.05	0.09	kua_Latn	0.11	0.11	0.11	0.08	**0.12**
khs_Latn	**0.14**	0.09	0.06	0.05	0.05	kub_Latn	0.09	**0.14**	0.05	0.05	0.05
khy_Latn	0.08	0.09	0.07	0.07	**0.14**	kud_Latn	0.07	**0.10**	0.06	0.05	0.05
khz_Latn	0.12	**0.16**	0.06	0.05	0.05	kue_Latn	0.07	**0.11**	0.06	0.05	0.07
kia_Latn	0.13	0.19	0.06	0.05	**0.23**	kuj_Latn	0.12	**0.12**	0.05	0.05	0.05
kij_Latn	0.07	**0.14**	0.07	0.05	0.06	kum_Cyrl	0.07	0.16	0.13	0.24	**0.45**
kik_Latn	0.14	**0.15**	0.05	0.05	0.05	kup_Latn	**0.18**	0.15	0.08	0.05	0.07
kin_Latn	0.14	0.13	0.14	0.06	**0.23**	kus_Latn	**0.12**	0.09	0.10	0.05	0.05
kir_Cyrl	0.07	0.20	**0.65**	0.65	0.61	kvg_Latn	**0.11**	0.09	0.06	0.05	0.06
kix_Latn	0.08	**0.12**	0.07	0.05	0.05	kvj_Latn	**0.17**	0.13	0.06	0.05	0.05
kjb_Latn	0.15	0.11	0.05	0.05	**0.23**	kvn_Latn	**0.12**	0.09	0.08	0.05	0.06
kje_Latn	0.09	**0.18**	0.06	0.05	0.06	kwd_Latn	**0.19**	0.13	0.09	0.05	0.12
kjh_Cyrl	0.07	0.18	0.11	0.17	**0.36**	kwf_Latn	**0.21**	0.17	0.09	0.07	0.16
kjs_Latn	**0.13**	0.10	0.07	0.05	0.05	kwi_Latn	0.11	**0.17**	0.09	0.05	0.09
kki_Latn	0.16	**0.17**	0.14	0.10	0.14	kwj_Latn	0.10	**0.12**	0.06	0.05	0.05
kkj_Latn	0.09	**0.16**	0.06	0.05	0.06	kxc_Ethi	0.07	**0.09**	0.07	0.05	0.05
kle_Deva	0.07	0.14	0.15	0.11	**0.19**	kxm_Thai	0.07	0.08	**0.14**	0.06	0.08
kln_Latn	0.10	0.10	0.05	0.05	**0.12**	kxw_Latn	0.06	**0.07**	0.06	0.05	0.05
klv_Latn	0.09	**0.14**	0.13	0.05	0.09	kyc_Latn	0.07	**0.11**	0.06	0.05	0.06
kma_Latn	**0.12**	0.08	0.05	0.05	0.05	kyf_Latn	0.09	**0.13**	0.05	0.05	0.05
kmd_Latn	0.10	**0.11**	0.06	0.05	0.09	kyg_Latn	0.08	**0.09**	0.06	0.05	0.05
kmg_Latn	0.08	**0.08**	0.05	0.05	0.05	kyq_Latn	0.10	**0.12**	0.07	0.05	0.05
kmh_Latn	0.07	**0.10**	0.05	0.05	0.05	kyu_Mymr	0.07	**0.09**	0.05	0.05	0.05
kmk_Latn	0.10	0.10	0.06	0.05	**0.14**	kyz_Latn	**0.17**	0.10	0.05	0.05	0.05
kmm_Latn	0.12	0.09	0.05	0.05	**0.19**	kze_Latn	0.08	**0.11**	0.04	0.05	0.06
kmo_Latn	**0.10**	0.09	0.05	0.06	0.06	kzf_Latn	0.12	**0.18**	0.10	0.06	0.15
kmr_Cyrl	0.07	0.09	0.07	0.05	**0.24**	lac_Latn	**0.16**	0.05	0.06	0.05	0.11
kms_Latn	**0.13**	0.08	0.04	0.05	0.07	lai_Latn	0.16	0.13	0.07	0.08	**0.19**
kmu_Latn	0.07	**0.17**	0.10	0.05	0.08	laj_Latn	0.10	**0.11**	0.07	0.06	0.09
kmy_Latn	**0.12**	0.08	0.05	0.05	0.05	lam_Latn	0.09	0.14	0.07	0.07	**0.16**
kne_Latn	**0.15**	0.13	0.12	0.04	0.09	lao_Laoo	0.07	0.05	0.58	**0.67**	0.61

Table 15: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
lap_Latn	0.14	**0.15**	0.06	0.05	0.08	mbb_Latn	0.11	**0.20**	0.10	0.05	0.10
las_Latn	0.09	**0.09**	0.05	0.05	0.05	mbc_Latn	0.12	**0.13**	0.05	0.05	0.05
lat_Latn	0.14	0.30	0.55	**0.62**	0.56	mbd_Latn	**0.13**	0.12	0.11	0.05	0.10
lav_Latn	0.08	0.34	**0.62**	0.55	0.52	mbf_Latn	0.07	0.31	0.49	**0.57**	0.56
law_Latn	0.09	**0.09**	0.06	0.05	0.09	mbh_Latn	0.15	**0.15**	0.07	0.05	0.09
lbk_Latn	0.12	0.10	0.09	0.05	**0.14**	mbi_Latn	0.13	**0.17**	0.08	0.05	0.06
lcm_Latn	0.16	**0.20**	0.05	0.06	0.15	mbj_Latn	**0.16**	0.14	0.08	0.05	0.06
lcp_Thai	0.07	**0.08**	0.06	0.05	0.05	mbl_Latn	0.07	**0.11**	0.05	0.05	0.05
ldi_Latn	0.14	0.12	0.07	0.05	**0.19**	mbs_Latn	0.11	0.12	0.17	0.13	**0.19**
lee_Latn	**0.08**	0.05	0.07	0.05	0.05	mbt_Latn	**0.14**	0.12	0.07	0.05	0.09
lef_Latn	0.05	**0.13**	0.06	0.05	0.05	mca_Latn	**0.16**	0.10	0.05	0.05	0.06
leh_Latn	0.09	0.14	0.08	0.07	**0.15**	mcb_Latn	0.07	**0.11**	0.05	0.05	0.06
lem_Latn	0.07	**0.09**	0.05	0.05	0.06	mcd_Latn	0.05	**0.09**	0.05	0.05	0.06
leu_Latn	0.12	**0.14**	0.05	0.05	0.07	mcf_Latn	0.07	**0.10**	0.06	0.05	0.05
lew_Latn	0.07	0.13	0.08	0.05	**0.16**	mck_Latn	0.13	**0.15**	0.11	0.06	0.15
lex_Latn	**0.13**	0.10	0.08	0.05	0.05	mcn_Latn	0.09	**0.10**	0.07	0.06	0.10
lgg_Latn	0.09	**0.19**	0.05	0.05	0.13	mco_Latn	0.05	0.09	0.05	0.05	**0.13**
lgl_Latn	**0.20**	0.14	0.06	0.06	0.12	mcp_Latn	**0.09**	0.05	0.05	0.05	0.05
lgm_Latn	**0.12**	0.11	0.06	0.06	0.09	mcq_Latn	0.07	**0.12**	0.08	0.05	0.05
lhi_Latn	0.09	**0.12**	0.05	0.05	0.10	mcu_Latn	0.10	**0.20**	0.07	0.05	0.06
lhm_Latn	**0.12**	0.08	0.05	0.05	0.05	mda_Latn	0.06	**0.07**	0.05	0.05	0.05
lhu_Latn	**0.09**	0.08	0.06	0.05	0.06	mdy_Ethi	0.07	0.09	0.05	0.05	**0.15**
lia_Latn	**0.18**	0.16	0.05	0.05	0.05	med_Latn	0.07	**0.09**	0.06	0.05	0.07
lid_Latn	**0.16**	0.09	0.08	0.05	0.06	mee_Latn	0.11	**0.12**	0.05	0.05	0.06
lif_Deva	0.07	0.07	0.10	0.05	**0.13**	mej_Latn	0.07	**0.11**	0.09	0.05	0.08
lin_Latn	0.12	0.10	0.08	0.04	**0.13**	mek_Latn	0.08	0.10	0.08	0.05	**0.14**
lip_Latn	0.08	**0.12**	0.06	0.05	0.07	men_Latn	0.11	**0.13**	0.05	0.05	0.05
lis_Lisu	0.07	**0.08**	0.05	0.05	0.06	meq_Latn	**0.10**	0.07	0.07	0.05	0.05
lit_Latn	0.07	0.29	0.56	**0.60**	0.54	met_Latn	**0.19**	0.11	0.05	0.05	0.06
ljp_Latn	0.07	0.29	0.33	0.30	**0.39**	meu_Latn	0.10	**0.14**	0.10	0.05	0.08
llg_Latn	0.07	0.09	**0.13**	0.05	0.07	mfe_Latn	0.09	0.15	0.15	0.05	**0.36**
lln_Latn	**0.10**	0.09	0.05	0.05	0.05	mfh_Latn	0.07	**0.07**	0.06	0.05	0.07
lmk_Latn	**0.14**	0.11	0.07	0.05	0.05	mfi_Latn	**0.15**	0.07	0.06	0.05	0.06
lmp_Latn	0.09	**0.12**	0.05	0.05	0.05	mfk_Latn	0.09	**0.16**	0.05	0.05	0.05
lnd_Latn	0.09	0.13	0.10	0.06	**0.15**	mfq_Latn	**0.08**	0.05	0.05	0.05	0.06
lob_Latn	0.07	**0.10**	0.05	0.05	0.04	mfy_Latn	0.11	**0.15**	0.07	0.05	0.06
loe_Latn	0.10	0.21	0.10	0.08	**0.23**	mfz_Latn	**0.13**	0.09	0.05	0.05	0.05
log_Latn	0.11	**0.11**	0.05	0.05	0.05	mgh_Latn	**0.13**	0.10	0.04	0.05	0.08
lok_Latn	**0.13**	0.12	0.05	0.05	0.05	mgo_Latn	**0.15**	0.05	0.05	0.05	0.05
lol_Latn	0.07	**0.09**	0.06	0.05	0.09	mgr_Latn	0.17	0.13	0.10	0.07	**0.21**
lom_Latn	**0.11**	0.07	0.05	0.05	0.05	mhi_Latn	0.12	**0.12**	0.08	0.05	0.06
loq_Latn	0.08	**0.13**	0.05	0.05	0.06	mhl_Latn	0.10	**0.10**	0.05	0.05	0.05
loz_Latn	0.18	0.14	0.06	0.05	**0.29**	mhr_Cyrl	0.07	0.17	0.10	0.05	**0.26**
lsi_Latn	**0.13**	0.08	0.05	0.05	0.05	mhx_Latn	0.11	**0.12**	0.05	0.05	0.05
lsm_Latn	0.11	**0.16**	0.08	0.07	0.08	mhy_Latn	0.12	0.20	0.21	0.15	**0.26**
ltz_Latn	0.15	0.34	0.22	0.20	**0.41**	mib_Latn	0.09	**0.13**	0.07	0.06	0.13
luc_Latn	0.07	0.09	**0.11**	0.05	0.05	mic_Latn	0.10	**0.13**	0.08	0.05	0.06
lug_Latn	0.07	0.13	0.08	0.05	**0.22**	mie_Latn	0.08	**0.17**	0.06	0.05	0.12
luo_Latn	0.12	0.12	0.05	0.05	**0.15**	mif_Latn	0.09	**0.09**	0.07	0.05	0.07
lus_Latn	**0.17**	0.14	0.10	0.05	0.09	mig_Latn	0.13	**0.19**	0.05	0.05	0.07
lwo_Latn	0.12	**0.12**	0.05	0.05	0.05	mih_Latn	0.08	**0.13**	0.04	0.05	0.07
lww_Latn	0.11	**0.12**	0.06	0.05	0.05	mil_Latn	0.10	**0.11**	0.05	0.05	0.06
lzh_Hani	0.07	0.24	0.54	0.50	**0.59**	mim_Latn	0.11	**0.15**	0.05	0.05	0.06
maa_Latn	0.13	**0.14**	0.05	0.05	0.05	min_Latn	0.08	0.19	0.27	0.26	**0.43**
mad_Latn	0.10	0.22	0.23	0.19	**0.40**	mio_Latn	0.09	0.08	**0.15**	0.07	0.14
maf_Latn	0.11	**0.18**	0.06	0.05	0.05	mip_Latn	0.06	0.10	0.05	0.05	**0.11**
mag_Deva	0.07	0.22	0.38	0.32	**0.49**	miq_Latn	0.09	**0.16**	0.05	0.05	0.08
mah_Latn	**0.16**	0.12	0.05	0.05	0.14	mir_Latn	0.06	0.09	0.06	0.05	**0.14**
mai_Deva	0.07	0.23	0.31	0.43	**0.65**	mit_Latn	0.06	0.09	0.07	0.06	**0.12**
maj_Latn	0.09	**0.09**	0.05	0.05	0.05	miy_Latn	0.07	**0.10**	0.05	0.05	0.08
mak_Latn	0.10	**0.18**	0.10	0.06	0.18	miz_Latn	0.09	**0.14**	0.05	0.05	0.05
mal_Mlym	0.07	**0.12**	0.07	0.05	0.06	mjc_Latn	0.13	**0.13**	0.05	0.05	0.07
mam_Latn	0.12	0.11	0.04	0.04	**0.25**	mjw_Latn	0.08	**0.09**	0.08	0.05	0.05
maq_Latn	0.12	**0.15**	0.05	0.06	0.05	mkd_Cyrl	0.07	0.47	**0.74**	0.70	0.67
mar_Deva	0.07	0.30	0.57	**0.61**	0.59	mkl_Latn	**0.11**	0.05	0.06	0.05	0.05
mas_Latn	0.07	**0.17**	0.09	0.06	0.04	mkn_Latn	0.07	0.23	0.28	0.35	**0.44**
mau_Latn	0.07	**0.08**	0.05	0.05	0.05	mks_Latn	0.10	**0.15**	0.05	0.05	0.05
mav_Latn	**0.14**	0.12	0.07	0.05	0.05	mlg_Latn	0.12	0.08	0.37	0.45	**0.46**
maw_Latn	**0.18**	0.11	0.05	0.05	0.05	mlh_Latn	0.10	**0.10**	0.05	0.05	0.05
maz_Latn	0.10	**0.15**	0.05	0.05	0.10	mlp_Latn	0.07	**0.20**	0.06	0.05	0.08

Table 16: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
mlt_Latn	0.11	0.16	0.05	0.06	**0.29**
mmn_Latn	0.17	0.19	0.18	0.21	**0.32**
mmo_Latn	**0.17**	0.09	0.09	0.05	0.05
mmx_Latn	**0.14**	0.11	0.05	0.05	0.06
mna_Latn	**0.11**	0.08	0.05	0.05	0.05
mnb_Latn	0.10	**0.17**	0.06	0.05	0.16
mnf_Latn	0.11	**0.13**	0.05	0.05	0.06
mnh_Latn	0.07	**0.17**	0.07	0.05	0.09
mnk_Latn	0.09	**0.17**	0.05	0.05	0.07
mnx_Latn	0.11	**0.15**	0.08	0.06	0.05
moa_Latn	**0.08**	0.04	0.06	0.05	0.05
moc_Latn	0.08	**0.13**	0.06	0.05	0.05
mog_Latn	0.16	0.20	0.13	0.07	**0.21**
mop_Latn	0.20	0.10	0.07	0.06	**0.27**
mor_Latn	**0.14**	0.11	0.05	0.05	0.05
mos_Latn	0.11	**0.11**	0.06	0.05	0.06
mox_Latn	0.12	**0.15**	0.07	0.05	0.05
mpg_Latn	**0.12**	0.09	0.05	0.05	0.05
mpm_Latn	0.04	**0.15**	0.05	0.05	0.05
mps_Latn	0.15	**0.16**	0.05	0.06	0.07
mpt_Latn	**0.13**	0.11	0.07	0.05	0.07
mpx_Latn	0.09	**0.10**	0.07	0.05	0.05
mqb_Latn	**0.11**	0.09	0.04	0.05	0.05
mqj_Latn	0.11	**0.18**	0.12	0.05	0.16
mqy_Latn	0.11	**0.16**	0.13	0.05	0.11
mri_Latn	0.16	0.09	0.09	0.05	**0.19**
mrw_Latn	0.09	0.19	0.10	0.14	**0.31**
msa_Latn	0.08	0.22	0.42	0.42	**0.52**
msb_Latn	0.12	0.21	0.28	0.24	**0.49**
mse_Latn	**0.12**	0.09	0.08	0.05	0.05
msk_Latn	0.09	0.14	0.09	0.10	**0.28**
msm_Latn	0.12	0.10	0.07	0.06	**0.21**
msy_Latn	0.07	**0.09**	0.06	0.05	0.06
mta_Latn	**0.12**	0.10	0.05	0.05	0.05
mtg_Latn	**0.11**	0.09	0.05	0.05	0.05
mti_Latn	0.14	0.14	0.08	0.08	**0.15**
mtj_Latn	0.08	**0.10**	0.08	0.05	0.06
mto_Latn	0.11	**0.14**	0.05	0.05	0.05
mtp_Latn	0.11	**0.12**	0.05	0.05	0.05
mua_Latn	**0.16**	0.10	0.05	0.05	0.06
mug_Latn	**0.13**	0.11	0.05	0.06	0.07
muh_Latn	0.12	**0.18**	0.15	0.05	0.05
mup_Deva	0.07	0.28	0.35	0.32	**0.49**
mur_Latn	**0.14**	0.12	0.05	0.05	0.08
mux_Latn	**0.12**	0.11	0.06	0.05	0.05
muy_Latn	**0.11**	0.07	0.05	0.05	0.05
mva_Latn	0.07	**0.15**	0.07	0.05	0.07
mvn_Latn	**0.12**	0.09	0.05	0.05	0.05
mvp_Latn	0.11	0.12	0.15	0.05	**0.22**
mwm_Latn	**0.12**	0.08	0.05	0.05	0.05
mwq_Latn	0.10	**0.10**	0.06	0.05	0.05
mwv_Latn	0.07	**0.14**	0.10	0.05	0.13
mww_Latn	**0.10**	0.06	0.05	0.05	0.05
mxb_Latn	0.09	**0.14**	0.05	0.05	0.06
mxp_Latn	0.10	**0.12**	0.05	0.05	0.06
mxq_Latn	0.09	0.06	0.05	0.05	**0.10**
mxt_Latn	**0.13**	0.12	0.04	0.05	0.07
mxv_Latn	0.10	**0.16**	0.05	0.05	0.16
mya_Mymr	0.07	0.26	0.42	**0.61**	0.51
myb_Latn	0.07	**0.13**	0.07	0.05	0.09
myk_Latn	0.07	**0.12**	0.05	0.05	0.07
myu_Latn	0.07	**0.12**	0.09	0.05	0.06
myv_Cyrl	0.07	0.08	0.08	0.05	**0.19**
myw_Latn	0.07	**0.15**	0.06	0.05	0.05
myx_Latn	0.10	**0.12**	0.04	0.05	0.10
myy_Latn	0.07	0.08	**0.09**	0.05	0.06
mza_Latn	0.10	**0.13**	0.06	0.05	0.05
mzh_Latn	0.08	0.19	0.08	0.05	**0.24**
mzk_Latn	0.14	**0.14**	0.08	0.06	0.07
mzl_Latn	**0.10**	0.09	0.06	0.05	0.05
mzm_Latn	0.09	**0.09**	0.05	0.05	0.05
mzw_Latn	0.05	**0.09**	0.05	0.05	0.06
nab_Latn	0.07	**0.14**	0.05	0.05	0.05
naf_Latn	0.07	**0.15**	0.05	0.05	0.06
nak_Latn	0.11	**0.12**	0.04	0.05	0.08
nan_Latn	**0.14**	0.11	0.05	0.05	0.06
naq_Latn	0.09	**0.10**	0.05	0.05	0.07
nas_Latn	0.07	0.09	**0.11**	0.05	0.09
nav_Latn	**0.19**	0.09	0.05	0.05	0.05
naw_Latn	0.08	**0.10**	0.05	0.05	0.05
nbc_Latn	0.09	**0.12**	0.06	0.05	0.07
nbe_Latn	**0.17**	0.12	0.06	0.06	0.07
nbl_Latn	0.09	0.13	0.15	0.21	**0.29**
nbu_Latn	**0.15**	0.09	0.05	0.05	0.05
nca_Latn	0.07	**0.11**	0.06	0.06	0.06
nch_Latn	0.10	**0.12**	0.07	0.05	0.06
ncj_Latn	**0.14**	0.10	0.05	0.05	0.07
ncl_Latn	0.10	0.09	0.06	0.09	**0.13**
ncq_Laoo	0.07	0.05	**0.11**	0.04	0.10
nct_Latn	**0.12**	0.09	0.06	0.05	0.06
ncu_Latn	0.06	**0.09**	0.05	0.05	0.05
ndc_Latn	0.07	0.15	0.10	0.07	**0.16**
nde_Latn	0.09	0.13	0.15	0.21	**0.29**
ndi_Latn	**0.11**	0.10	0.06	0.05	0.05
ndj_Latn	**0.13**	0.11	0.06	0.05	0.12
ndo_Latn	0.11	0.11	0.09	0.05	**0.16**
ndp_Latn	0.10	**0.11**	0.10	0.05	0.07
nds_Latn	0.15	0.19	0.14	0.07	**0.27**
ndy_Latn	0.07	**0.14**	0.07	0.06	0.14
ndz_Latn	0.09	**0.15**	0.05	0.05	0.05
neb_Latn	**0.12**	0.07	0.05	0.05	0.05
nep_Deva	0.07	0.32	0.62	0.64	**0.68**
nfa_Latn	0.07	**0.09**	0.06	0.05	0.05
nfr_Latn	**0.15**	0.11	0.07	0.05	0.05
ngc_Latn	0.11	**0.14**	0.07	0.05	0.14
ngp_Latn	0.13	0.17	0.16	0.12	**0.19**
ngu_Latn	0.06	0.09	0.05	0.06	**0.15**
nhd_Latn	0.12	**0.17**	0.09	0.05	0.10
nhe_Latn	0.10	**0.13**	0.07	0.05	0.08
nhg_Latn	0.10	0.12	0.05	0.05	**0.14**
nhi_Latn	**0.12**	0.10	0.06	0.05	0.08
nho_Latn	0.16	**0.17**	0.07	0.05	0.12
nhr_Latn	**0.17**	0.14	0.05	0.05	0.07
nhu_Latn	**0.16**	0.10	0.05	0.05	0.05
nhw_Latn	0.08	**0.14**	0.07	0.05	0.06
nhx_Latn	0.13	0.14	0.08	0.05	**0.19**
nhy_Latn	0.14	**0.16**	0.05	0.06	0.15
nii_Latn	**0.14**	0.09	0.05	0.05	0.05
nij_Latn	0.09	**0.23**	0.18	0.16	0.23
nim_Latn	0.07	**0.12**	0.06	0.05	0.06
nin_Latn	0.07	**0.13**	0.08	0.05	0.07
niq_Latn	0.09	**0.10**	0.05	0.05	0.07
niy_Latn	**0.11**	0.05	0.08	0.05	0.05
njb_Latn	**0.17**	0.13	0.05	0.05	0.05
njm_Latn	**0.16**	0.09	0.06	0.05	0.06
njn_Latn	0.09	**0.12**	0.05	0.05	0.05
njo_Latn	**0.12**	0.11	0.05	0.05	0.06
njz_Latn	0.08	**0.13**	0.05	0.05	0.05
nkf_Latn	0.13	**0.16**	0.06	0.05	0.06
nki_Latn	0.10	0.13	0.05	0.05	**0.26**
nko_Latn	0.10	**0.10**	0.05	0.05	0.05
nlc_Latn	0.11	**0.12**	0.05	0.05	0.05
nld_Latn	0.28	0.43	**0.60**	0.58	0.53
nlg_Latn	0.20	**0.21**	0.07	0.09	0.21
nma_Latn	0.07	**0.12**	0.08	0.05	0.05
nmf_Latn	0.08	**0.12**	0.05	0.05	0.06
nmh_Latn	0.09	**0.10**	0.05	0.06	0.06
nmo_Latn	0.10	**0.10**	0.06	0.05	0.06
nmz_Latn	**0.15**	0.12	0.08	0.05	0.10
nnb_Latn	0.10	**0.14**	0.07	0.05	0.10

Table 17: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
nng_Latn	0.07	**0.09**	0.07	0.05	0.06	oym_Latn	0.07	**0.12**	0.05	0.05	0.05
nnh_Latn	0.08	**0.14**	0.07	0.05	0.08	ozm_Latn	**0.13**	0.06	0.06	0.05	0.05
nnl_Latn	0.12	**0.12**	0.07	0.05	0.06	pab_Latn	**0.12**	0.05	0.05	0.05	0.05
nno_Latn	0.15	0.46	**0.58**	0.56	0.43	pad_Latn	0.13	**0.15**	0.06	0.05	0.06
nnp_Latn	0.07	**0.08**	0.07	0.05	0.05	pag_Latn	0.14	0.14	0.20	0.17	**0.33**
nnq_Latn	0.14	**0.15**	0.11	0.10	0.14	pah_Latn	0.09	**0.15**	0.06	0.05	0.05
nnw_Latn	**0.07**	0.05	0.05	0.05	0.05	pam_Latn	0.13	0.18	0.11	0.11	**0.38**
noa_Latn	0.07	**0.08**	0.05	0.06	0.05	pan_Guru	0.07	0.31	0.58	0.67	**0.69**
nob_Latn	0.16	0.38	0.59	**0.60**	0.56	pao_Latn	0.10	**0.13**	0.07	0.05	0.08
nod_Thai	0.07	0.09	0.47	0.50	**0.50**	pap_Latn	0.15	0.31	0.30	0.23	**0.52**
nog_Cyrl	0.07	0.16	0.18	0.38	**0.41**	pau_Latn	0.16	0.18	0.06	0.05	**0.21**
nop_Latn	0.09	**0.15**	0.05	0.05	0.05	pbb_Latn	**0.17**	0.12	0.07	0.05	0.07
nor_Latn	0.16	0.38	**0.60**	0.60	0.55	pbc_Latn	**0.17**	0.12	0.05	0.05	0.05
not_Latn	0.07	0.09	**0.13**	0.06	0.11	pbi_Latn	**0.13**	0.06	0.05	0.05	0.07
nou_Latn	**0.16**	0.11	0.11	0.06	0.13	pbl_Latn	0.10	0.16	0.13	0.05	**0.26**
nph_Latn	0.08	**0.10**	0.09	0.05	0.05	pck_Latn	0.12	0.14	0.06	0.05	**0.19**
npi_Deva	0.07	0.32	0.59	0.66	**0.67**	pcm_Latn	0.19	0.18	0.30	0.29	**0.45**
npl_Latn	0.10	0.09	0.05	0.07	**0.18**	pdc_Latn	0.19	0.14	0.14	0.15	**0.27**
npo_Latn	**0.13**	0.09	0.07	0.05	0.05	pdt_Latn	0.17	0.18	0.17	0.12	**0.34**
npy_Latn	0.09	**0.13**	0.11	0.05	0.07	pes_Arab	0.07	0.42	**0.66**	0.66	0.63
nre_Latn	0.10	**0.15**	0.07	0.05	0.07	pez_Latn	0.08	**0.23**	0.09	0.05	0.10
nri_Latn	0.11	**0.12**	0.09	0.05	0.09	pfe_Latn	**0.10**	0.05	0.05	0.05	0.05
nsa_Latn	0.07	**0.12**	0.09	0.05	0.06	pib_Latn	0.07	**0.11**	0.04	0.05	0.06
nse_Latn	0.12	0.17	0.13	0.07	**0.23**	pio_Latn	0.07	0.09	0.06	0.05	**0.12**
nsm_Latn	**0.13**	0.07	0.06	0.05	0.06	pir_Latn	0.10	**0.11**	0.06	0.05	0.05
nsn_Latn	**0.15**	0.09	0.06	0.07	0.12	pis_Latn	**0.21**	0.11	0.12	0.06	0.20
nso_Latn	0.11	0.13	0.12	0.05	**0.27**	pjt_Latn	0.07	**0.09**	0.05	0.05	0.08
nst_Latn	**0.18**	0.10	0.05	0.05	0.06	pkb_Latn	0.11	0.15	0.12	0.07	**0.28**
nsu_Latn	**0.13**	0.10	0.06	0.05	0.12	plg_Latn	**0.16**	0.13	0.08	0.05	0.08
ntp_Latn	0.07	**0.10**	0.05	0.05	0.04	pls_Latn	0.07	0.19	0.07	0.14	**0.27**
ntr_Latn	0.07	**0.12**	0.05	0.05	0.05	plt_Latn	0.12	0.05	0.38	**0.54**	0.50
ntu_Latn	0.07	**0.08**	0.06	0.05	0.05	plu_Latn	**0.13**	0.08	0.05	0.05	0.05
nuj_Latn	0.11	**0.14**	0.06	0.05	0.07	plw_Latn	0.14	**0.19**	0.10	0.06	0.19
nus_Latn	**0.13**	0.10	0.05	0.05	0.05	pma_Latn	0.14	**0.16**	0.07	0.05	0.06
nuy_Latn	**0.23**	0.10	0.05	0.05	0.05	pmf_Latn	0.11	**0.22**	0.10	0.09	0.20
nvm_Latn	0.07	**0.11**	0.05	0.05	0.05	pmx_Latn	**0.09**	0.08	0.06	0.06	0.06
nwb_Latn	**0.14**	0.06	0.05	0.05	0.05	pne_Latn	0.08	**0.23**	0.09	0.05	0.11
nwi_Latn	**0.15**	0.13	0.05	0.05	0.07	pny_Latn	**0.08**	0.05	0.05	0.05	0.05
nwx_Deva	0.07	0.16	0.18	0.14	**0.29**	poe_Latn	0.13	**0.13**	0.05	0.05	0.06
nxd_Latn	0.07	**0.09**	0.07	0.05	0.07	poh_Latn	0.11	0.09	0.12	0.05	**0.37**
nya_Latn	0.07	0.14	0.08	0.06	**0.26**	poi_Latn	0.12	**0.15**	0.05	0.07	0.12
nyf_Latn	0.15	0.19	0.21	0.17	**0.25**	pol_Latn	0.09	0.48	0.60	**0.65**	0.61
nyn_Latn	0.09	0.11	0.06	0.05	**0.20**	pon_Latn	0.14	**0.21**	0.08	0.05	0.08
nyo_Latn	0.07	**0.16**	0.05	0.05	0.15	por_Latn	0.16	0.52	0.57	**0.64**	0.61
nyy_Latn	0.11	**0.16**	0.08	0.05	0.09	pos_Latn	0.12	0.17	0.06	0.06	**0.27**
nza_Latn	0.07	**0.10**	0.05	0.05	0.05	poy_Latn	0.14	**0.18**	0.08	0.05	0.07
nzi_Latn	0.09	**0.16**	0.05	0.05	0.05	ppk_Latn	0.15	0.15	0.06	0.04	**0.16**
nzm_Latn	**0.11**	0.09	0.08	0.06	0.06	ppo_Latn	0.10	**0.18**	0.05	0.05	0.05
obo_Latn	**0.15**	0.12	0.05	0.05	0.07	pps_Latn	0.10	**0.11**	0.06	0.05	0.08
ojb_Cans	0.07	**0.12**	0.05	0.05	0.06	prf_Latn	0.12	0.20	0.15	0.13	**0.26**
oji_Latn	**0.11**	0.09	0.05	0.05	0.07	pri_Latn	0.07	**0.10**	0.05	0.05	0.05
ojs_Latn	0.07	**0.08**	0.05	0.05	0.06	prk_Latn	0.09	**0.13**	0.06	0.05	0.10
oku_Latn	**0.12**	0.11	0.05	0.05	0.05	prq_Latn	0.07	**0.08**	0.05	0.05	0.05
okv_Latn	0.13	**0.22**	0.14	0.08	0.13	prs_Arab	0.07	0.43	**0.66**	0.64	0.64
old_Latn	**0.13**	0.09	0.08	0.06	0.06	pse_Latn	0.07	0.28	0.36	0.38	**0.39**
omb_Latn	**0.17**	0.16	0.10	0.06	0.06	pss_Latn	0.10	**0.13**	0.06	0.05	0.08
omw_Latn	0.07	**0.08**	0.05	0.05	0.05	ptp_Latn	0.10	**0.11**	0.05	0.05	0.05
ong_Latn	0.07	**0.17**	0.07	0.05	0.06	ptu_Latn	0.11	0.15	0.14	0.05	**0.20**
ons_Latn	**0.11**	0.09	0.05	0.05	0.05	pua_Latn	0.08	0.09	0.09	0.05	**0.15**
ood_Latn	**0.16**	0.11	0.05	0.05	0.05	pui_Latn	0.09	**0.14**	0.05	0.06	0.06
opm_Latn	0.07	**0.14**	0.07	0.05	0.05	pwg_Latn	**0.18**	0.14	0.06	0.08	0.12
ori_Orya	0.07	0.04	0.58	**0.75**	0.65	pww_Thai	0.07	0.08	**0.10**	0.05	0.05
ory_Orya	0.07	0.04	0.56	**0.75**	0.64	pxm_Latn	0.08	**0.14**	0.06	0.05	0.05
oss_Cyrl	0.07	0.10	0.07	0.05	**0.11**	qub_Latn	0.08	0.12	0.06	0.06	**0.17**
otd_Latn	0.07	**0.25**	0.12	0.11	0.14	quc_Latn	0.18	0.14	0.07	0.05	**0.37**
ote_Latn	**0.08**	0.07	0.05	0.05	0.06	quf_Latn	0.07	**0.10**	0.05	0.05	0.06
otm_Latn	**0.10**	0.08	0.05	0.05	0.05	qug_Latn	0.07	0.11	0.09	0.05	**0.12**
otn_Latn	0.09	**0.11**	0.05	0.05	0.05	quh_Latn	0.07	0.12	0.07	0.05	**0.30**
otq_Latn	**0.14**	0.08	0.06	0.05	0.06	qul_Latn	0.07	0.14	0.06	0.07	**0.32**
ots_Latn	**0.11**	0.10	0.05	0.05	0.10	qup_Latn	0.07	**0.13**	0.05	0.05	0.13

Table 18: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
quw_Latn	0.07	0.10	0.07	0.05	**0.18**	shp_Latn	0.07	**0.12**	0.06	0.05	0.05
quy_Latn	0.07	0.11	0.07	0.06	**0.27**	shu_Latn	0.09	**0.20**	0.16	0.11	0.19
quz_Latn	0.07	0.10	0.07	0.05	**0.24**	sig_Latn	**0.13**	0.08	0.05	0.05	0.05
qva_Latn	0.07	0.10	0.07	0.05	**0.18**	sil_Latn	**0.14**	0.07	0.05	0.05	0.05
qvc_Latn	0.09	**0.11**	0.06	0.05	0.05	sim_Latn	0.08	**0.10**	0.06	0.05	0.07
qve_Latn	0.09	0.13	0.06	0.05	**0.33**	sin_Sinh	0.07	0.16	0.51	**0.67**	0.57
qvh_Latn	0.12	0.12	0.05	0.07	**0.24**	sja_Latn	0.10	**0.10**	0.05	0.05	0.05
qvi_Latn	0.06	**0.12**	0.06	0.05	0.10	sld_Latn	**0.14**	0.10	0.05	0.05	0.05
qvm_Latn	0.07	0.13	0.06	0.05	**0.19**	slk_Latn	0.09	0.48	**0.69**	0.64	0.56
qvn_Latn	0.07	0.10	0.05	0.06	**0.14**	sll_Latn	0.07	**0.11**	0.07	0.05	0.08
qvo_Latn	0.10	**0.11**	0.06	0.05	0.08	slv_Latn	0.17	0.50	**0.63**	0.60	0.60
qvs_Latn	0.09	0.10	0.05	0.05	**0.18**	sme_Latn	0.15	**0.17**	0.09	0.05	0.14
qvw_Latn	0.09	0.10	0.05	0.05	**0.13**	smk_Latn	0.10	0.10	0.08	0.06	**0.27**
qvz_Latn	0.09	0.10	0.06	0.05	**0.13**	sml_Latn	0.13	0.12	0.17	0.10	**0.23**
qwh_Latn	0.06	0.14	0.09	0.05	**0.22**	smo_Latn	0.10	0.07	0.08	0.05	**0.29**
qxh_Latn	0.07	0.11	0.04	0.05	**0.15**	smt_Latn	0.11	0.15	0.05	0.05	**0.21**
qxl_Latn	0.07	**0.11**	0.07	0.05	0.08	sna_Latn	0.07	0.11	0.11	0.08	**0.18**
qxn_Latn	0.07	0.15	0.07	0.05	**0.23**	snc_Latn	**0.15**	0.12	0.05	0.05	0.06
qxo_Latn	0.09	0.11	0.05	0.06	**0.23**	snd_Arab	0.07	0.19	0.61	**0.67**	0.61
qxr_Latn	0.07	0.13	0.10	0.05	**0.14**	snf_Latn	**0.14**	0.11	0.06	0.05	0.06
rad_Latn	0.09	**0.09**	0.06	0.05	0.06	snn_Latn	0.14	**0.17**	0.09	0.05	0.05
rai_Latn	0.16	**0.18**	0.05	0.07	0.12	snp_Latn	**0.12**	0.11	0.06	0.05	0.09
rap_Latn	0.13	0.13	0.06	0.05	**0.21**	snw_Latn	0.09	**0.11**	0.05	0.05	0.05
rar_Latn	0.10	0.07	0.06	0.05	**0.22**	sny_Latn	0.07	**0.13**	0.06	0.05	0.08
rav_Deva	0.07	0.09	**0.17**	0.05	0.07	som_Latn	0.08	0.09	0.31	0.39	**0.43**
raw_Latn	0.12	**0.14**	0.05	0.05	0.06	sop_Latn	0.15	0.14	0.07	0.05	**0.20**
rej_Latn	0.12	0.25	0.20	0.18	**0.31**	soq_Latn	**0.19**	0.17	0.05	0.07	0.08
rel_Latn	**0.15**	0.12	0.08	0.05	0.06	sot_Latn	0.13	0.10	0.09	0.05	**0.18**
rgu_Latn	0.07	0.07	0.04	0.04	**0.15**	soy_Latn	**0.16**	0.07	0.05	0.05	0.05
ria_Latn	0.08	**0.10**	0.06	0.05	0.06	spa_Latn	0.11	0.49	0.64	**0.69**	0.58
rim_Latn	0.13	**0.16**	0.05	0.06	0.07	spl_Latn	0.07	**0.12**	0.05	0.05	0.05
rjs_Deva	0.07	0.13	0.26	0.22	**0.28**	spp_Latn	**0.10**	0.08	0.06	0.05	0.09
rkb_Latn	**0.12**	0.07	0.05	0.05	0.08	sps_Latn	0.14	**0.17**	0.05	0.05	0.05
rmc_Latn	0.12	0.17	0.17	0.09	**0.18**	spy_Latn	0.07	**0.09**	0.05	0.05	0.07
rmo_Latn	**0.17**	0.16	0.08	0.06	0.11	sqi_Latn	0.10	0.33	**0.68**	0.66	0.65
rmy_Latn	0.12	**0.23**	0.10	0.06	0.22	sri_Latn	0.07	**0.13**	0.04	0.05	0.06
rnl_Latn	0.11	**0.14**	0.05	0.05	0.09	srm_Latn	0.12	0.09	0.06	0.05	**0.21**
ron_Latn	0.11	0.50	0.62	**0.65**	0.53	srn_Latn	0.07	0.15	0.07	0.05	**0.42**
roo_Latn	0.07	**0.10**	0.05	0.05	0.05	srp_Latn	0.09	0.47	0.59	0.59	**0.63**
rop_Latn	0.20	**0.20**	0.06	0.05	0.20	srq_Latn	**0.16**	0.07	0.11	0.07	0.10
row_Latn	0.07	**0.08**	0.06	0.05	0.08	ssd_Latn	0.12	**0.17**	0.05	0.05	0.05
rro_Latn	0.08	**0.11**	0.07	0.05	0.05	ssg_Latn	**0.13**	0.06	0.11	0.06	0.06
rub_Latn	0.13	**0.13**	0.08	0.05	0.08	ssw_Latn	0.07	0.11	0.09	0.12	**0.24**
ruf_Latn	0.14	**0.20**	0.10	0.09	0.11	ssx_Latn	0.11	**0.13**	0.07	0.05	0.06
rug_Latn	0.10	**0.13**	0.06	0.05	0.06	stn_Latn	**0.19**	0.16	0.11	0.05	0.15
run_Latn	0.16	0.15	0.09	0.06	**0.27**	stp_Latn	**0.09**	0.04	0.05	0.05	0.05
rus_Cyrl	0.07	0.50	0.55	**0.67**	0.64	sua_Latn	**0.18**	0.13	0.05	0.05	0.05
rwo_Latn	0.07	**0.10**	0.07	0.06	0.05	suc_Latn	**0.13**	0.11	0.06	0.05	0.08
sab_Latn	0.07	**0.10**	0.08	0.05	0.06	sue_Latn	0.13	**0.14**	0.08	0.05	0.06
sag_Latn	0.11	0.19	0.10	0.06	**0.20**	suk_Latn	**0.16**	0.13	0.07	0.07	0.09
sah_Cyrl	0.07	0.12	0.08	0.05	**0.30**	sun_Latn	0.09	0.33	0.45	**0.50**	0.45
saj_Latn	0.05	**0.10**	0.05	0.05	0.08	sur_Latn	**0.15**	0.11	0.06	0.05	0.10
san_Taml	**0.07**	0.05	0.07	0.05	0.05	sus_Latn	0.12	**0.15**	0.04	0.05	0.05
sas_Latn	0.11	0.22	0.28	0.24	**0.30**	suz_Deva	0.07	0.10	0.11	0.06	**0.27**
sat_Latn	**0.12**	0.08	0.06	0.05	0.06	swe_Latn	0.13	0.48	**0.73**	0.60	0.59
sba_Latn	**0.12**	0.11	0.06	0.05	0.11	swg_Latn	0.21	0.27	0.25	0.34	**0.35**
sbd_Latn	**0.12**	0.09	0.06	0.06	0.05	swh_Latn	0.12	0.31	0.50	**0.57**	0.54
sbl_Latn	0.12	0.08	0.18	0.12	**0.21**	swk_Latn	0.11	0.13	0.04	0.06	**0.19**
sck_Deva	0.07	0.17	0.28	0.44	**0.47**	swp_Latn	0.08	**0.10**	0.08	0.06	0.06
sda_Latn	0.11	**0.16**	0.09	0.05	0.13	sxb_Latn	0.10	0.13	0.08	0.05	**0.14**
sdq_Latn	0.06	0.15	0.12	0.10	**0.16**	sxn_Latn	0.07	0.09	0.05	0.05	**0.18**
seh_Latn	0.13	0.11	0.07	0.06	**0.23**	syb_Latn	**0.13**	0.09	0.10	0.05	0.11
ses_Latn	**0.14**	0.09	0.07	0.05	0.07	syc_Syrc	0.07	0.05	0.05	0.08	**0.10**
sey_Latn	0.06	**0.10**	0.05	0.05	0.05	syl_Latn	**0.07**	0.06	0.05	0.05	0.05
sgb_Latn	0.14	0.22	0.17	0.10	**0.31**	szb_Latn	0.07	**0.21**	0.04	0.05	0.06
sgw_Ethi	0.07	0.09	0.10	0.13	**0.24**	tab_Cyrl	0.07	0.11	**0.12**	0.05	0.10
sgz_Latn	0.07	**0.13**	0.06	0.05	0.07	tac_Latn	0.12	**0.20**	0.05	0.05	0.07
shi_Latn	**0.13**	0.07	0.05	0.05	0.07	taj_Deva	0.07	0.13	0.14	0.09	**0.20**
shk_Latn	**0.11**	0.07	0.06	0.05	0.07	tam_Taml	0.07	0.35	0.53	0.56	**0.60**
shn_Mymr	**0.07**	0.05	0.06	0.05	0.05	tap_Latn	0.14	0.18	0.10	0.08	**0.20**

Table 19: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m	lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
taq_Latn	0.10	**0.11**	0.07	0.05	0.06	tro_Latn	**0.15**	0.12	0.07	0.05	0.07
tar_Latn	0.10	**0.10**	0.05	0.05	0.05	trp_Latn	**0.10**	0.08	0.06	0.05	0.05
tat_Cyrl	0.07	0.31	0.12	0.15	**0.45**	trq_Latn	0.05	**0.12**	0.05	0.05	0.07
tav_Latn	**0.13**	0.11	0.05	0.05	0.09	trs_Latn	0.06	**0.10**	0.07	0.05	0.10
taw_Latn	**0.14**	0.09	0.07	0.05	0.07	tsg_Latn	0.11	0.17	0.15	0.11	**0.27**
tbc_Latn	0.09	**0.12**	0.05	0.05	0.06	tsn_Latn	0.12	0.12	0.09	0.05	**0.23**
tbg_Latn	0.07	**0.14**	0.08	0.05	0.06	tsw_Latn	0.07	**0.12**	0.07	0.05	0.08
tbk_Latn	0.07	0.17	0.11	0.11	**0.27**	tsz_Latn	0.08	0.10	0.08	0.05	**0.14**
tbl_Latn	0.12	**0.12**	0.12	0.05	0.06	ttc_Latn	0.14	**0.20**	0.10	0.05	0.09
tbo_Latn	0.12	**0.13**	0.10	0.05	0.05	tte_Latn	0.07	0.07	**0.08**	0.05	0.05
tbw_Latn	0.11	0.15	0.08	0.06	**0.25**	ttq_Latn	0.09	0.09	0.07	0.06	**0.10**
tby_Latn	**0.14**	0.12	0.06	0.05	0.12	ttr_Cyrl	0.07	0.31	0.18	0.13	**0.42**
tbz_Latn	0.07	**0.09**	0.05	0.05	0.05	tuc_Latn	**0.18**	0.10	0.05	0.05	0.05
tca_Latn	0.07	**0.07**	0.05	0.05	0.07	tue_Latn	0.07	**0.10**	0.04	0.05	0.05
tcc_Latn	0.09	**0.10**	0.05	0.05	0.05	tuf_Latn	0.11	**0.13**	0.10	0.05	0.06
tcs_Latn	0.21	0.19	0.11	0.06	**0.21**	tui_Latn	**0.17**	0.14	0.08	0.05	0.07
tcz_Latn	**0.12**	0.11	0.09	0.05	0.05	tuk_Latn	0.11	0.11	0.22	0.22	**0.44**
tdt_Latn	0.15	0.15	0.09	0.05	**0.36**	tul_Latn	0.12	**0.18**	0.05	0.05	0.05
ted_Latn	**0.10**	0.09	0.05	0.05	0.05	tum_Latn	0.13	**0.22**	0.10	0.07	0.21
tee_Latn	0.06	0.07	0.06	0.05	**0.14**	tuo_Latn	**0.12**	0.09	0.04	0.05	0.08
tel_Telu	0.07	0.30	0.60	0.67	**0.67**	tur_Latn	0.11	0.29	**0.68**	0.68	0.63
tem_Latn	**0.12**	0.05	0.06	0.05	0.05	tvk_Latn	0.11	**0.19**	0.08	0.05	0.10
teo_Latn	0.09	**0.12**	0.05	0.07	0.08	twb_Latn	0.10	**0.12**	0.05	0.05	0.06
ter_Latn	0.12	**0.13**	0.06	0.05	0.06	twi_Latn	0.10	**0.15**	0.05	0.05	0.13
tet_Latn	0.07	0.11	0.05	0.05	**0.13**	twu_Latn	0.12	0.15	**0.16**	0.05	0.07
tfr_Latn	0.12	**0.14**	0.08	0.05	0.05	txq_Latn	0.07	**0.15**	0.09	0.05	0.06
tgk_Cyrl	0.07	0.19	0.05	0.04	**0.31**	txu_Latn	0.13	**0.17**	0.07	0.05	0.05
tgl_Latn	0.13	0.29	0.47	0.55	**0.55**	tyv_Cyrl	0.07	0.12	0.19	0.18	**0.44**
tgo_Latn	0.09	**0.14**	0.05	0.05	0.05	tzh_Latn	0.08	0.10	0.09	0.05	**0.22**
tgp_Latn	0.15	**0.21**	0.08	0.09	0.09	tzj_Latn	0.13	0.15	0.09	0.06	**0.21**
tha_Thai	0.07	0.08	0.56	**0.60**	0.56	tzo_Latn	0.08	0.11	0.07	0.05	**0.30**
thk_Latn	**0.16**	0.10	0.04	0.05	0.05	ubr_Latn	**0.15**	0.13	0.06	0.05	0.10
thl_Deva	0.07	0.24	0.34	0.44	**0.45**	ubu_Latn	**0.13**	0.07	0.07	0.05	0.06
tif_Latn	0.07	**0.10**	0.05	0.05	0.08	udm_Cyrl	0.07	0.10	0.07	0.05	**0.20**
tih_Latn	0.09	0.11	0.09	0.05	**0.26**	udu_Latn	**0.19**	0.11	0.05	0.05	0.08
tik_Latn	**0.09**	0.07	0.05	0.05	0.05	uig_Cyrl	0.07	0.20	0.13	0.14	**0.44**
tim_Latn	0.07	**0.11**	0.06	0.05	0.06	ukr_Cyrl	0.07	0.40	0.64	**0.67**	0.57
tir_Ethi	0.07	0.06	0.27	0.22	**0.38**	upv_Latn	0.10	**0.12**	0.06	0.05	0.05
tiy_Latn	0.15	**0.17**	0.08	0.06	0.08	ura_Latn	0.07	**0.08**	0.05	0.05	0.05
tke_Latn	0.13	**0.14**	0.06	0.05	0.09	urb_Latn	**0.14**	0.11	0.12	0.05	0.05
tku_Latn	0.10	0.09	0.06	0.05	**0.15**	urd_Arab	0.07	0.37	0.49	**0.67**	0.56
tlb_Latn	0.09	**0.13**	0.07	0.05	0.09	urk_Thai	0.07	**0.09**	0.07	0.05	0.05
tlf_Latn	0.07	0.07	**0.09**	0.05	0.08	urt_Latn	0.06	**0.13**	0.08	0.05	0.06
tlh_Latn	0.22	**0.29**	0.24	0.13	0.29	ury_Latn	**0.14**	0.10	0.05	0.05	0.06
tlj_Latn	**0.19**	0.14	0.11	0.05	0.12	usa_Latn	0.07	**0.10**	0.06	0.05	0.05
tmc_Latn	0.10	**0.12**	0.05	0.05	0.08	usp_Latn	0.18	0.11	0.07	0.05	**0.24**
tmd_Latn	0.07	**0.08**	0.05	0.05	0.05	uth_Latn	0.07	**0.10**	0.09	0.05	0.07
tna_Latn	0.11	0.12	**0.13**	0.05	0.07	uvh_Latn	0.07	**0.09**	0.07	0.05	0.05
tnk_Latn	0.11	**0.11**	0.05	0.05	0.04	uvl_Latn	0.09	**0.16**	0.06	0.05	0.09
tnn_Latn	**0.13**	0.10	0.07	0.05	0.07	uzb_Latn	0.09	0.14	0.54	**0.59**	0.58
tnp_Latn	**0.12**	0.07	0.05	0.07	0.06	uzn_Cyrl	0.07	0.14	0.07	0.10	**0.47**
tnr_Latn	**0.13**	0.07	0.05	0.05	0.06	vag_Latn	0.10	**0.11**	0.05	0.05	0.06
tob_Latn	0.07	**0.12**	0.04	0.05	0.09	vap_Latn	**0.19**	0.12	0.06	0.05	0.17
toc_Latn	0.06	**0.09**	0.05	0.05	0.05	var_Latn	0.10	**0.13**	0.07	0.05	0.06
toh_Latn	0.11	0.12	0.06	0.06	**0.22**	ven_Latn	0.11	**0.12**	0.06	0.05	0.11
toi_Latn	0.07	0.13	0.08	0.06	**0.24**	vid_Latn	0.11	**0.14**	0.11	0.09	0.09
toj_Latn	0.12	0.06	0.07	0.05	**0.29**	vie_Latn	0.09	0.38	0.54	**0.63**	0.53
ton_Latn	0.09	0.08	0.05	0.05	**0.26**	viv_Latn	0.07	**0.11**	0.06	0.05	0.05
too_Latn	0.10	**0.11**	0.06	0.05	0.11	vmy_Latn	**0.13**	0.10	0.05	0.05	0.10
top_Latn	0.08	0.13	0.05	0.05	**0.17**	vun_Latn	**0.13**	0.10	0.06	0.05	0.05
tos_Latn	0.06	**0.07**	0.05	0.05	0.07	vut_Latn	**0.08**	0.05	0.05	0.05	0.05
tpi_Latn	0.17	0.17	0.09	0.06	**0.31**	waj_Latn	**0.10**	0.08	0.06	0.05	0.06
tpm_Latn	**0.14**	0.12	0.06	0.05	0.06	wal_Latn	**0.15**	0.10	0.06	0.06	0.13
tpp_Latn	0.13	**0.15**	0.06	0.05	0.10	wap_Latn	0.11	**0.11**	0.06	0.05	0.06
tpt_Latn	0.14	0.07	0.09	0.05	**0.15**	war_Latn	0.11	0.16	0.15	0.14	**0.37**
tpz_Latn	**0.12**	0.11	0.06	0.05	0.06	way_Latn	0.10	**0.12**	0.07	0.05	0.05
tqb_Latn	0.07	**0.11**	0.08	0.05	0.05	wba_Latn	0.09	0.10	0.08	0.06	**0.11**
tqo_Latn	**0.12**	0.08	0.06	0.05	0.05	wbm_Latn	0.09	**0.13**	0.06	0.05	0.09
trc_Latn	0.05	**0.14**	0.05	0.05	0.07	wbp_Latn	0.07	**0.07**	0.06	0.05	0.05
trn_Latn	0.12	**0.15**	0.06	0.06	0.05	wca_Latn	0.07	**0.14**	0.05	0.05	0.08

Table 20: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
wer_Latn	0.09	**0.15**	0.05	0.05	0.05
whk_Latn	0.11	**0.17**	0.07	0.05	0.11
wim_Latn	0.07	**0.08**	0.06	0.05	0.08
wiu_Latn	0.12	**0.13**	0.05	0.06	0.05
wmw_Latn	0.14	0.16	0.23	0.31	**0.41**
wnc_Latn	0.07	**0.12**	0.07	0.06	0.05
wnu_Latn	0.11	**0.13**	0.05	0.05	0.05
wob_Latn	**0.11**	0.06	0.05	0.05	0.05
wol_Latn	**0.16**	0.12	0.07	0.05	0.07
wos_Latn	**0.16**	0.10	0.08	0.05	0.06
wrs_Latn	**0.15**	0.10	0.06	0.05	0.05
wsg_Telu	0.07	0.09	**0.13**	0.08	0.07
wsk_Latn	0.12	**0.15**	0.08	0.05	0.10
wuv_Latn	**0.18**	0.09	0.09	0.05	0.06
wwa_Latn	**0.16**	0.08	0.05	0.06	0.05
xal_Cyrl	0.07	0.12	0.08	0.05	**0.14**
xav_Latn	0.11	**0.13**	0.08	0.05	0.10
xbr_Latn	0.09	**0.09**	0.08	0.05	0.07
xed_Latn	**0.11**	0.10	0.06	0.05	0.07
xho_Latn	0.09	0.14	0.21	0.30	**0.34**
xla_Latn	**0.13**	0.08	0.08	0.05	0.05
xmm_Latn	0.14	0.30	**0.42**	0.40	0.40
xnn_Latn	0.07	0.11	0.10	0.08	**0.19**
xog_Latn	0.07	0.16	0.06	0.06	**0.22**
xon_Latn	0.06	**0.17**	0.05	0.05	0.05
xpe_Latn	0.08	**0.11**	0.05	0.05	0.06
xrb_Latn	0.11	**0.11**	0.05	0.05	0.05
xsb_Latn	0.11	0.14	0.11	0.08	**0.23**
xsi_Latn	0.09	**0.13**	0.05	0.05	0.05
xsm_Latn	**0.19**	0.08	0.05	0.05	0.05
xsr_Deva	0.07	**0.09**	0.05	0.05	0.06
xsu_Latn	0.13	**0.15**	0.05	0.05	0.08
xtd_Latn	0.14	**0.16**	0.05	0.05	0.07
xtm_Latn	0.07	**0.15**	0.06	0.06	0.08
xtn_Latn	0.09	**0.16**	0.07	0.06	0.13
xuo_Latn	**0.10**	0.08	0.05	0.05	0.05
yaa_Latn	0.07	**0.11**	0.06	0.05	0.06
yad_Latn	**0.11**	0.09	0.05	0.05	0.05
yal_Latn	**0.15**	0.13	0.06	0.05	0.07
yam_Latn	**0.13**	0.05	0.05	0.05	0.05
yan_Latn	0.10	**0.13**	0.05	0.05	0.05
yao_Latn	0.13	0.13	0.06	0.05	**0.15**
yap_Latn	0.13	0.14	0.07	0.05	**0.22**
yaq_Latn	0.16	**0.16**	0.07	0.05	0.06
yas_Latn	**0.13**	0.10	0.05	0.05	0.05
yat_Latn	**0.11**	0.05	0.05	0.05	0.06
yaz_Latn	0.07	**0.12**	0.08	0.05	0.05
ybb_Latn	0.07	**0.09**	0.05	0.05	0.05
yby_Latn	0.07	**0.08**	0.07	0.07	0.05
ycn_Latn	**0.10**	0.09	0.05	0.05	0.05
yim_Latn	**0.13**	0.12	0.09	0.05	0.06
yka_Latn	0.09	0.14	0.10	0.07	**0.26**
yle_Latn	0.07	**0.13**	0.05	0.05	0.05
yli_Latn	0.11	**0.17**	0.09	0.05	0.10
yml_Latn	0.08	**0.08**	0.05	0.05	0.06
yom_Latn	0.09	0.16	0.06	0.05	**0.21**
yon_Latn	**0.12**	0.11	0.11	0.05	0.09
yor_Latn	0.11	**0.14**	0.10	0.05	0.10
yrb_Latn	**0.19**	0.10	0.11	0.05	0.06
yre_Latn	0.08	**0.11**	0.05	0.05	0.05
yss_Latn	0.10	**0.12**	0.08	0.05	0.08
yua_Latn	0.16	**0.16**	0.11	0.05	0.13
yue_Hani	0.07	0.40	**0.60**	0.60	0.56
yuj_Latn	**0.14**	0.08	0.09	0.06	0.07
yut_Latn	0.11	**0.14**	0.05	0.05	0.05
yuw_Latn	0.10	**0.12**	0.09	0.05	0.05
yuz_Latn	0.07	**0.12**	0.10	0.05	0.10
yva_Latn	0.13	**0.15**	0.06	0.05	0.06
zaa_Latn	0.10	0.20	0.20	0.07	**0.29**
zab_Latn	0.07	0.08	0.13	0.07	**0.16**

lan_script	BOW	mBert	XLM-R-B	XLM-R-L	Glot500-m
zac_Latn	0.12	**0.20**	0.09	0.09	0.18
zad_Latn	**0.15**	0.10	0.04	0.05	0.05
zae_Latn	**0.14**	0.13	0.10	0.05	0.06
zai_Latn	0.08	0.21	0.13	0.09	**0.25**
zam_Latn	0.09	**0.16**	0.07	0.05	0.13
zao_Latn	**0.14**	0.09	0.06	0.05	0.06
zar_Latn	0.11	**0.17**	0.06	0.05	0.08
zas_Latn	0.07	**0.16**	0.07	0.06	0.13
zat_Latn	0.13	0.11	0.11	0.06	**0.13**
zav_Latn	**0.07**	0.06	0.05	0.05	0.06
zaw_Latn	0.07	0.06	0.06	0.05	**0.07**
zca_Latn	0.21	0.14	0.18	0.06	**0.21**
zho_Hani	0.07	0.39	**0.63**	0.63	0.59
zia_Latn	**0.14**	0.11	0.06	0.05	0.06
ziw_Latn	0.13	0.17	0.14	0.11	**0.23**
zlm_Latn	0.07	0.47	0.68	**0.71**	0.62
zoc_Latn	0.11	0.08	0.06	0.05	**0.11**
zom_Latn	0.10	0.16	0.13	0.05	**0.27**
zos_Latn	0.15	**0.16**	0.05	0.06	0.14
zpc_Latn	**0.13**	0.12	0.11	0.05	0.12
zpi_Latn	0.13	**0.16**	0.09	0.05	0.08
zpl_Latn	0.07	0.13	0.13	0.06	**0.17**
zpm_Latn	**0.17**	0.14	0.05	0.06	0.08
zpo_Latn	0.10	**0.15**	0.13	0.06	0.10
zpq_Latn	0.07	**0.10**	0.06	0.05	0.09
zpt_Latn	0.11	0.11	0.10	0.05	**0.16**
zpu_Latn	**0.14**	0.08	0.05	0.05	0.06
zpv_Latn	**0.10**	0.08	0.05	0.05	0.05
zpz_Latn	0.05	0.07	**0.08**	0.05	0.05
zsm_Latn	0.07	0.53	**0.71**	0.63	0.58
zsr_Latn	0.09	**0.12**	0.07	0.05	0.09
ztq_Latn	0.10	0.13	0.10	0.08	**0.19**
zty_Latn	0.11	0.06	0.09	0.05	**0.12**
zul_Latn	0.07	0.11	0.23	0.33	**0.37**
zyb_Latn	**0.15**	0.10	0.06	0.05	0.05
zyp_Latn	0.10	**0.15**	0.05	0.05	0.06

Table 21: zero-shot score of BOW, mBERT, XLM-R-B, XLM-R-L, and Glot500-m.