Identifying Open Challenges in Language Identification

Rob van der Goot
IT University of Copenhagen
robv@itu.dk

Abstract

Automatic language identification is a core problem of many Natural Language Processing (NLP) pipelines. A wide variety of architectures and benchmarks have been proposed with often near-perfect performance. Although previous studies have focused on certain challenging setups (i.e. cross-domain, short inputs), a systematic comparison is missing. We propose a benchmark that allows us to test for the effect of input size, training data size, domain, number of languages, scripts, and language families on performance. We evaluate five popular models on this benchmark and identify which open challenges remain for this task as well as which architectures achieve robust performance. We find that cross-domain setups are the most challenging (although arguably most relevant), and that number of languages, variety in scripts, and variety in language families have only a small impact on performance. We also contribute practical takeaways: training with 1,000 instances per language and a maximum input length of 100 characters is enough for robust language identification. Based on our findings, we train an accurate (94.41%) multi-domain language identification model on 2,034 languages, for which we also provide an analysis of the remaining errors.[1]

1 Introduction

Language identification is a crucial step for many Natural Language Processing (NLP) pipelines. It can for example be used to provide conditional information for multi-lingual models (Conneau and Lample, 2019), decide which model to use, to pre-filter raw data for training language models (Kreutzer et al., 2022), or even for filtering data to annotate. High performances have been obtained with numerous benchmarks and models, although previous work has also focused on specific challenging dimensions, for example, social media data (Lui and Baldwin, 2014a), in cross-domain setups (Lui and Baldwin, 2011), or for short input texts (Toftrup et al., 2021).

To the best of our knowledge, evaluation of language identification has been fragmented and setups vary across many dimensions, including domains, datasets, metrics, number of languages, input size, amount of training data, number of scripts, and number of language families. In this paper, we will carefully compose a benchmark consisting of a variety of open sources, and we constrain our setup among several dimensions to identify open challenges in language identification and perform an in-depth robustness evaluation of common language identification models. We inspect the effect of size of input per instance, number of instances per language, number of supported languages, domain overlap, language family distance, and script overlap. We use a variety of types of popular language identification tools, which we re-train for all evaluations for a fair comparison. Our contributions are:

- We provide a dataset for language identification that allows for creating subsets that are stratified among a variety of dimensions.
- We evaluate five commonly used language identification models on a variety of benchmarks sampled from our main collection of datasets.
- We identify open problems for the task of language classification of text.
- We release a language classification model with the largest set of supported languages (2,034) to date, and analyze remaining errors.

2 Models

We focus on the most popular and best-performing methods for language identification and pick a com-

[1] Code and best-performing models are available on: https://bitbucket.org/robvanderg/langid_problems

Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 18207–18227
July 27 - August 1, 2025 ©2025 Association for Computational Linguistics

monly used implementation for each method. For a more complete overview of architectures used for language identification, we refer to Jauhiainen et al. (2019). For all our experiments we train the models from scratch for a fair comparison. We ended up with the following categories and models:

N-gram overlap Character n-gram frequencies provide a footprint of a text, and have indeed for a long time been used successfully as features in machine-learning text classification problems. Note that character n-gram overlaps are more robust as compared to words n-grams, as for many languages it is unclear what the word boundaries are. We use the approach proposed by Cavnar et al. (1994) as implemented by van Noord (1997), named **Textcat**. In short, the model first builds profiles for languages by looking for the 400 most frequent 1-5 character n-grams, and then calculates the distance of this ranked list to a ranked list of 1-5 character n-grams of the target utterance. The input text will be classified to the most similar profile.

Naive Bayes A variety of machine learning algorithms, such as Naive Bayes (e.g. Zampieri et al., 2014, 2015), SVM (e.g. Majliš, 2012), and Logistic Regression (e.g. Bhargava et al., 2015; Camposampiero et al., 2022) have successfully been used for the language identification task. Character n-grams are commonly used as robust input features. We choose a Naïve Bayes (**NB**) classifier since it has shown to perform well (Brown, 2014) and is computationally more efficient (especially with large amounts of classes). Initially, we used `langid.py`, a widely used standalone Python package for language identification. However, performance were unexpectedly poor, perhaps because of our conversion python3. Hence, we reimplemented a multinomial Naive Bayes with default settings of scikit-learn (Pedregosa et al., 2011) based on character n-grams for which we did a small hyperparameter search on a setup with 100 languages (and 100 instances of 100 characters). We ended up with 1-5 grams, binary feature representations (as opposed to counts or TF-IDF), and 100,000 maximum features.

This specific implementation uses Naïve Bayes with byte n-grams (1-4) as input features.

Static embeddings **FastText** (Joulin et al., 2017) is a popular toolkit for language identification. It uses a bag of character n-grams as input with a single hidden layer to obtain predictions. The original fastText model for language identification has been trained on 176 languages (Wikipedia data), but follow-up work has shown strong performance in a variety of setups (e.g. Burchell et al., 2023; Kargaran et al., 2023a). We follow the hyperparameter setup of Burchell et al. (2023) and Kargaran et al. (2023a), except the number of epochs (we use 10 instead of 2), because we use less data.

LSTM LSTMs (Hochreiter and Schmidhuber, 1997; Graves and Schmidhuber, 2005) were first used for language identification by Cazamias et al. (2015). More recently, Bi-LSTM models have been shown to perform competitively, especially for short input texts (Toftrup et al., 2021). As we could not find a good reference for generic hyperparameters, we did a small fine-tuning search detailed in Appendix A.

Contextualized language models Just like for many other NLP tasks, transformer-based language models have shown promising results for language identification. For example, in the recent Discriminating Between Similar Languages (DSL-TL) shared task (Aepli et al., 2023), the highest performances were obtained by language models. One downside of language models for this task is the constrained subword vocabulary, certain scripts, and languages that are underrepresented will be harder to classify, and in some cases with only unknown characters even impossible. We evaluated a range of multi-lingual language models on a subset of our data, and will do our main experiments with **Glot500** (ImaniGooghari et al., 2023), due to its superior performance (comparison to other language models is available in Appendix C). We did not experiment with generative language models, as previous work (Robinson et al., 2023; Chen et al., 2023) has shown subpar performance while they are computationally expensive.

3 Data

3.1 Sources

We selected data sources based on the quality of the language labels, diversity in domains/languages and availability. We removed parts of datasets that had multiple domain labels, were taken from other datasets in the list, or were translated Wikipedia data. We limit our studies to mono-lingual utterances, as opposed to language identification on code-switched data (Doğruöz et al., 2021; Winata et al., 2023). The data sources we list are all pub-

Dataset	langs	scripts	fams	domains
MIL-TALE	2,110	47	139	wiki, political, religious, grammar
UDHR	397	38	61	rights
OpenLID	139	25	16	literature, news, wiki, social, grammar, subtitles, spoken
MassiveSumm	77	24	13	news
TwitUser	59	20	13	social
UD	54	11	17	medical, news, academic, wiki, legal, nonfiction, learner-essays, fiction, social, grammar-examples, reviews, religious, spoken
Total	2,176/ 7,850	51/ 163	145/ 298	

Table 1: Dataset statistics. The languages counts (langs) are based on the ISO-639-3 standard, script counts on ISO 15924, and language families (fams) on Glottolog (Hammarström et al., 2023).

licly available and not tokenized. They are also not (consistently) sentence-segmented, so in some cases, the input is a sentence, and in other cases, they are paragraphs (or tweets in the case of Twitter). We removed all classes with less than 2,000 instances, so that we can have 1,000 instances for evaluation, and still have enough data left to train on. Further label cleaning is described in Appendix D. Dataset statistics are provided in Table 1.

MIL-TALE To the best of our knowledge, MIL-TALE (Brown, 2014) is the publicly available dataset with the widest language coverage. At the time of downloading (15-11-2023), it contained data for 2,221 languages from which 2,110 were left after our only keeping languages with more than 2,000 utterances. A vast majority of the data is of religious nature.

Universal Declaration of Human Rights A small (~90 lines) standardized text that is translated to many languages is the "Universal Declaration of Human Rights" (UDHR). We use this only as test data and scraped the most recent collection from `http://unicode.org/udhr/d/`. We used this dataset because of its distinct domain, and wide coverage of languages.

OpenLID Open-LID (Burchell et al., 2023) is a collection of already existing datasets (including MIL-TALE, which we already included, so we exclude the Open-LID version). We use the version of the dataset without sampling.

MassiveSumm Varab and Schluter (2021) automatically collected summarization data from the news domain with the clear desiderata of inclusivity and language variety. We use the original texts as input text (not the summaries). It should be noted that the utterances in this dataset are quite long (they are paragraphs), but we usually use only the first 100 characters for our experiments.[2]

Twituser A smaller dataset, with a focus on a single domain, is created by (Lui and Baldwin, 2014b). They collect Twitter messages based on language identification on the set of their tweets. Once they are certain that a user is tweeting mainly monolingual, they sample a small amount (to avoid user bias) of each user. The size of this dataset is relatively small, but we include it because it can give us valuable insights into the performance of language classification on less standard (web) domains.

UD We include Universal Dependency treebanks from UD v2.12 (Zeman et al., 2023). This dataset is rich in domain varieties, and we expect the language labels to be mostly accurate[3] since each dataset is manually annotated for syntax by speakers of the language. To ensure accurate domain information, we exclude all multi-domain treebanks.

3.2 Exploratory Data Analysis

We list basic dataset statistics in Table 1. These statistics highlight the variety of properties of the datasets, where some have a more distinctive domain (rights/social), others are rich in domain varieties (UD, OpenLID), and MIL-TALE for example has a much larger amount of languages covered. The total coverage of languages is now close to ~50%, as out of the ~7,000 languages that exist, only ~4,000 are estimated to have a "developed writing system".[4] For the scripts, we cover only

[2] Because we show that there is little effect of using longer inputs in Section 4.1

[3] We do not take into account the word level labels

[4] `https://web.archive.org/web/20230113104023/https://www.ethnologue.com/enterprise-faq/`

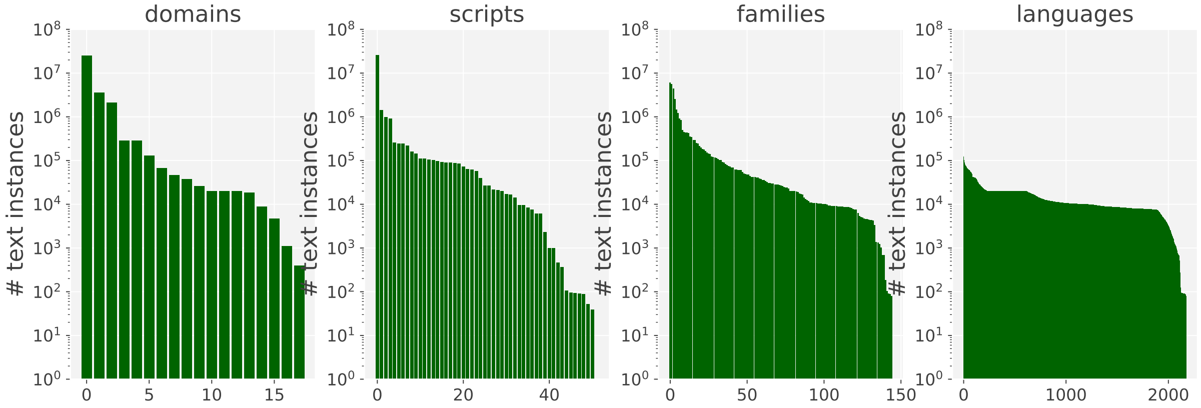


Figure 1: The number of text instances belonging to each category in our data among multiple dimensions.

about 1/3, and for language families the coverage is again ~50% (the 23 isolates in our data are not included in this count).

We also plot the number of datasets as well as utterances per domain, script, family, and language to get an overview of the amount of variety/information available across these dimensions. Figure 1 shows the distribution of text instances across different dimensions: domains, scripts, families, and languages. All of these have a similar shape, although for the dimensions with more labels, the largest set is smaller. They also show a long tail of low frequent labels (note that the y-axis is log-scale), though we did filter out the smallest sets.

4 Experiments

We carefully design our data splits for each of the dimensions that we will evaluate, hence we split this section into the different dimensions. We will start with data size, which includes the number of languages, amount of utterances per language, and length per utterance (Section 4.1). After this, we choose appropriate values of these dimensions and evaluate the other dimensions: language families (Section 4.2), scripts (Section 4.3), and domains (Section 4.4). Finally, we train a model on all our data and evaluate the effect of merging writing script with language labels (Section 4.5). For all our experiments we sample 1,000 utterances per language to use as test data[5] and report the average scores over three seeds. For ease of interpretation and because our evaluation sets are always balanced we use accuracy for all our main evaluations (i.e. unless mentioned otherwise).

4.1 Amount of information

Setup We first vary the amount of information given to the model, we vary among three dimensions 1) a maximum of 10, 100, or 1,000 characters of each instance, which we choose to (very approximately) match, words, sentences and paragraphs (it should be noted that some of our data sources are sentence-split, so they are not much longer than 100 characters) 2) 10, 100, or 1,000 utterances per language 3) sets of 10, 100, or 1,000 languages. We use different samples of languages for each random seed. We do not control for language family, script, or domain here, so the data is multi-domain (but we pick one random domain and script per language). Note that we have thus trained a total of 81 models per architecture (3 character sizes * 3 utterance sizes * 3 amount of languages * 3 seeds) for this experiment.

Results Figure 2 shows all results for the different amounts of information among all dimensions. In general, performance is more dependent on the setup as opposed to the choice of the model. GLOT500 achieves robust performance amongst most settings; this is most likely due to the fact that it is the only pre-trained model (for other models there are no publicly massively multilingual pre-trained representations), and the only clear breakdown happens at 10 utterances or characters per language. It is clear that the other models are more dependent on the amount of information, they have much stronger gains when having more utterances per language. Having access to 100 characters performs on-par for 1,000 characters for most setups, showing that there is enough signal in lengths that approximate match sentences. Our results also

how-many-languages-world-are-unwritten-0

[5] we do not tune any models, so we do not use a development split

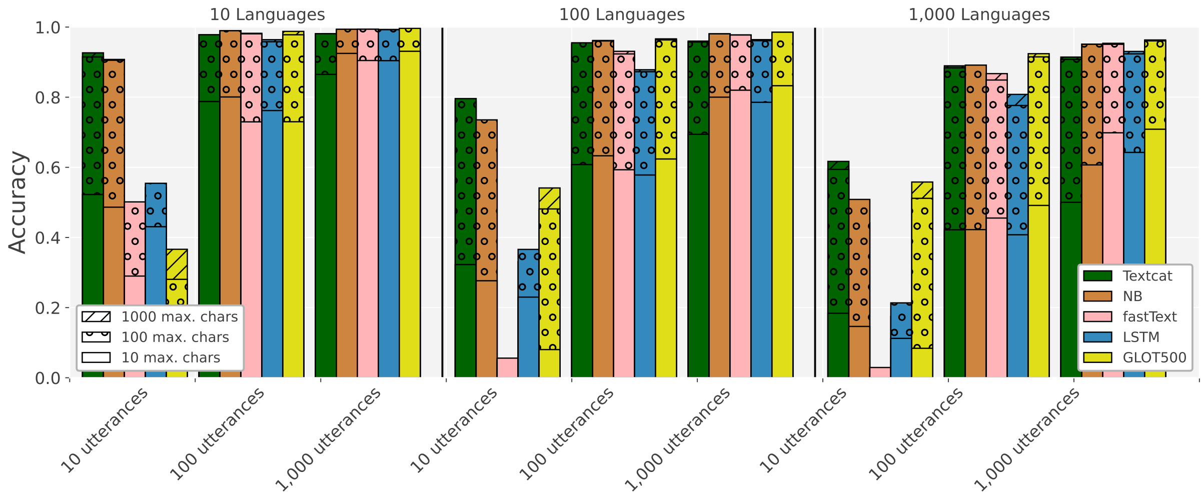
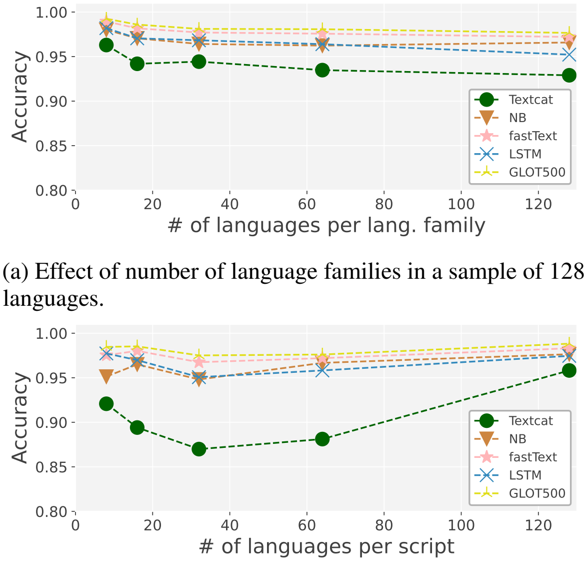


Figure 2: Comparison of all models when considering different amounts of information.



(a) Effect of number of language families in a sample of 128 languages.

(b) Effect of number of scripts in a sample of ~128 languages (note that it was impossible to sample this exact amount for the smaller numbers of scripts).

Figure 3: Performance with varying amounts of language families/scripts in a sample of ~128 languages. Concretely, this means that at the right of the figures, we have just a single language family or script, whereas on the left we have 8 languages for 16 language families/scripts.

show that 1,000 utterances with 100 characters lead to a very high performance across all numbers of languages, which has practical implications both for training time and requirements for low-resource languages. Standard deviation for all settings is reported in Appendix B and shows that the variance is mainly dependent on the setup, where setups with few languages or few characters per instance have a larger variance.

4.2 Number of language families

Setup We evaluate the effect of the number of language families in a setup by sampling a preset number of languages for each language family, where we have a total of 128 languages. On one extreme, we can have a single language family with 128 languages, on the other extreme, we can have 16 language families with each 8 languages (note that 128 families with 1 language is not possible with our dataset, as we only have 51 language families). We hypothesized that classifying within a language family is more challenging, as the languages are more likely to be similar. Similar to our other settings, we run for three seeds, with different sets of languages (and language families).

Results The results (Figure 3a) show that there is a small negative effect when the number of languages within a language family increases for almost all models. Textcat and the LSTM show a slightly higher sensitivity to an increasing amount of languages within a family, but drops overall are marginal.

4.3 Number of scripts

Setup We follow a similar setup as for language families, but group the languages by their main script. It should be noted that it was not possible for the settings with 8 and 16 scripts per language to obtain 126 total languages, so here we were limited to respectively 72 and 64 languages.

Results The number of scripts (Figure 3b) has a slightly different trend compared to language fam-

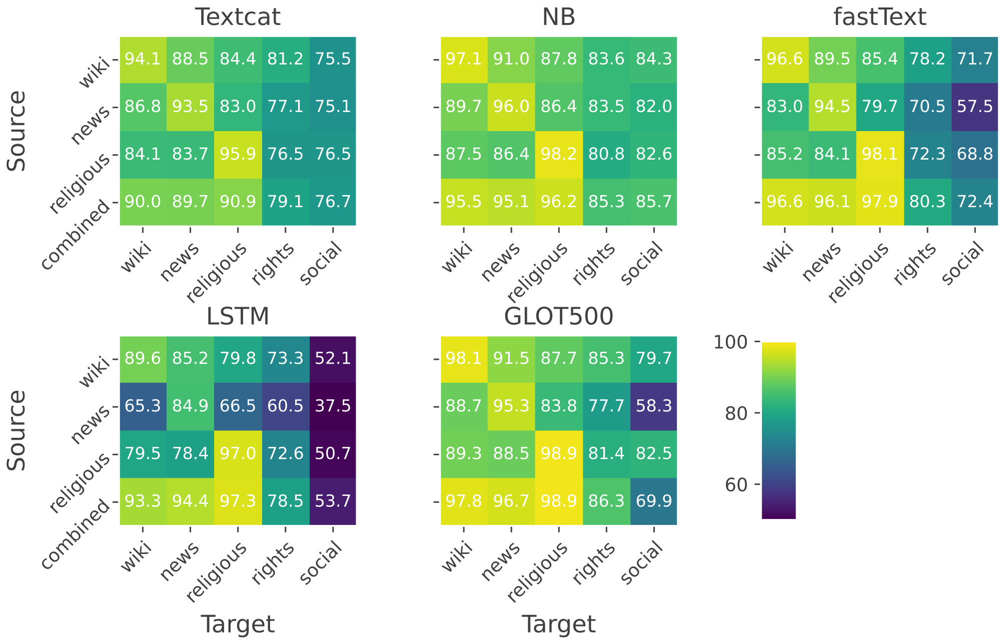


Figure 4: Accuracies in cross-domain setups (the diagonals are in-domain).

ily (Figure 3a), and has a larger effect on performance. There is a drop in performance when having 32 families per script, but when the number increases performance increases again. We hypothesize that this is due to the fact that the models can learn more accurate representations of their features due to a higher overlap. For example, there will be more overlap for character n-grams for fastText, and when there are more occurrences of a single feature, the model can learn a better representation for this feature.

4.4 Domain effect

Setup To evaluate the effect of domain transfer, we train single domain models on the largest three domains (to ensure a large number of languages and enough data), and evaluate the performance across these domains, as well as on the 'rights' and the 'social' domain. We also train a model on the three source domains jointly, to evaluate whether a multi-domain model is more robust against domain shift. The rights domain consists of the data from the Universal Declaration of Human Rights (UDHR). The 'social' domain consists mostly of Twituser data. We use these two domains, as they are not included in training (because the amount of data per language is small), but they have a wide language coverage.

Results The confusion matrices (Figure 4) show that the models have a different sensitivity to domain shift. In general, the losses in performance when going to cross-domain settings are large, showing that this is one of the main open challenges for this task. Textcat shows to be remarkably robust, outperforming fastText and LSTM for all test-only domains in all setups (last two columns). NB shows to be the most robust model with respect to domain shift, performing well (>80) for all settings. As opposed to previous settings, GLOT500 is not the most robust model and shows performance drops especially on the test-only domains. The multi-domain model ('Combined' in Figure 4) performs better for most model/target domain combinations, an outlier being GLOT500 for social media data. Especially for the more computationally demanding models, performance of some in-domain settings even improved, which is most likely an effect of dataset size. However, this finding indicates that multi-domain training should be considered when building robust language classifiers.

4.5 A language classifier for ~2,000 languages

Setup In the initial size experiments (Section 4.1), we saw that performance between sets of 100 and 1,000 languages only became marginally worse. Therefore, we experiment with scaling up the number of languages further. We still use 1,000 utterances per language for evaluation as well as for training, so exclude all languages with less than 2,000 utterances, resulting in a set of 2,034 lan-

Training	1,000	2,034		2,075	
Testing	1,000	1,000	2,034	1,000	2,075
Textcat	90.78$_{\pm.11}$	88.05$_{\pm.44}$	87.99$_{\pm.01}$	88.72$_{\pm.47}$	88.66$_{\pm.01}$
NB	95.10$_{\pm.23}$	92.44$_{\pm.46}$	92.29$_{\pm.00}$	93.00$_{\pm.48}$	92.81$_{\pm.01}$
fastText	95.09$_{\pm.04}$	93.43$_{\pm.12}$	93.41$_{\pm.02}$	93.46$_{\pm.16}$	93.42$_{\pm.01}$
LSTM	92.36$_{\pm.26}$	90.88$_{\pm.47}$	90.83$_{\pm.06}$	90.92$_{\pm.47}$	90.72$_{\pm.07}$
GLOT500	96.07$_{\pm.21}$	94.37$_{\pm.30}$	94.41$_{\pm.01}$	94.38$_{\pm.32}$	94.39$_{\pm.02}$

Table 2: Results (accuracy + standard deviation) on the test splits for a single classifier for 1,000 languages, 2,034 languages, and 2,075 language-script combinations. We evaluate all of them on the 1,000 language sample to gauge the effect of the number of languages.

guages. For all previous experiments, we used a single sub-dataset for each language, meaning that it would be in a single script. For this experiment, we train across domains and datasets, and will thus end up with having multiple scripts per language. To evaluate the effect of this, we also train a model that predicts the language code plus the script (as a single label), similar as in previous work (Brown, 2014; Team et al., 2022; Kargaran et al., 2023a), which leads to a total of 2,075 labels. We also check performance on the 1,000 samples of each seed as a control setting.

Results Results of the evaluation on the 2,034 and 2,075 languages compared to the 1,000 language setting from Section 4.1 are reported in Table 2. There is a consistent drop for all models of approximately 1.5-2.5 percentage points when going from 1,000 languages to 2,034, which is surprisingly large as the drop in performance is similar when going from 100 to 1,000 languages (Section 4.1), which is a factor 10 increase in size (instead of a factor 2). Performance for the 2,034 languages are very close to the 2,075 language-script combinations. This confirms similar contemporary findings by Agarwal et al. (2025), who find that language classification for languages with different scripts performs well as long as all scripts are included during training. The evaluation on the 1,000 languages show that performance of the larger (> 2,000) models is not affected as compared to the 1,000 language model.

5 Analysis

To provide additional insights beyond accuracy, we look into the computational complexity (Section 5.1), and provide a more detailed error analysis of the 2,037 languages model (Section 5.2).

Model	# params
Textcat	40,000
NB	100,000
fastText	4,434,860
LSTM	15,158,772
GLOT500	395,687,155

Table 3: Number of learned weights (# params) per model when training on 100 characters, 1,000 instances, and 100 languages.

5.1 Computational complexity

Setup Instead of evaluating the run-time or carbon emissions of our models, we opt for comparing the total number of weights in a model. We found that there is a too large discrepancy in the efficiency of the implementation of our models, [6] and prioritize an architectural comparison over an implementation comparison. Therefore, we use the number of weights learned in a model as a proxy to model complexity. We inspect the "average" models, trained on 100 characters, 1,000 instances, and 100 languages for these sizes.

Results The number of parameters (Table 3) show a substantial diversity, where the largest model is 10,000 times larger than the smaller model. If we compare the results of the models on the same setup, we see that Textcat obtains 94.5% of the performance of GLOT500, with only 0.01 % of the weights, which hence might be the preferable option.

[6] Concretely, fastText and GLOT500 seem to be optimized very extensively, while Textcat is extremely slow, although it is computationally the simplest.

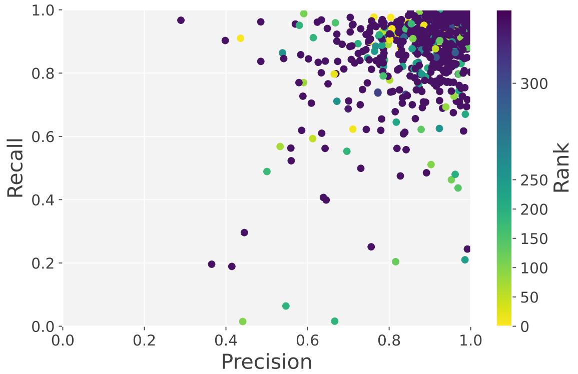


Figure 5: Precision and recall, each circle represents a language. The Wikipedia size is represented by color, where yellow (bright) indicates a high rank (i.e. large size).

5.2 Analysis of largest model

Subwords Even the best-performing model GLOT500 (with 1,000 utterances and 100 characters) still has an error rate of 5.59% on when evaluating our most inclusive setup with 2,034 languages (Section 4.5). Our first hunch was that the model encounters unknown characters, and therefore can not represent the input. However, the training data only had 0.184% of subwords represented as the special UNK token. There were two outliers (Vai and Yintale) with unknown subword rates of 45.7% and 42.8%, followed by 5 languages with a % between 10-20, and only 12 languages with a % between 1-10. Perhaps surprisingly, these languages with more unknown subwords were not among the worst-performing languages. This can be explained by the fact that these are low-resource languages. Even if a part of the text is converted to UNK tokens, the remaining part is distinctive enough for an accurate classification (because their scripts also often are uncommon).

Precision and recall We plot the precision against the recall to find the main weaknesses. Figure 5 shows that precision is generally higher than recall, and even for the most challenging languages, precision is generally above 40%. Arguably, this (precision > recall) is desirable when language identification is used as a filtering step, as we do not want false positives to dilute our data. We also rank the languages based on their Wikipedia size[7], to check whether better-represented languages are more likely to be overestimated (low precision, high recall), but can find no such trend in the results. We were expecting mainly English to have issues with data purity and overrepresentiveness, which is indeed confirmed by a precision of 82.75 and a recall of 87.80, although it should be well represented in the GLOT500 training data.

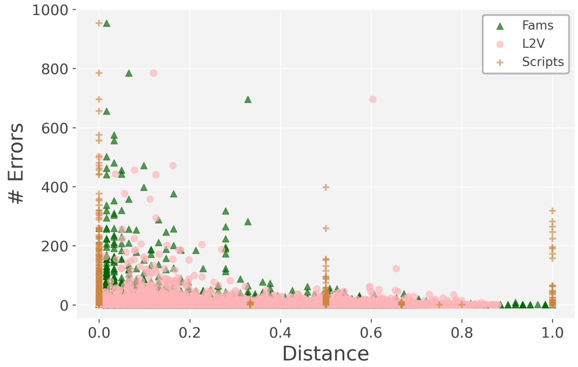


Figure 6: Number of error counts for language pairs plotted against a variety of distance metrics. We normalized all distance metrics to be between 0-1.

Most common confusions From all the errors of the model, 96.7% are cases where both the predicted and the gold language label are used within the same script, and 67.1% are within the same language family. Based on these results, we will perform a more in-depth study into investigating which features correlate with error counts of language pairs. The features we use are: cosine distance over lang2vec representations (Littell et al., 2017)[8], ratio of script overlap (because the other 2 metrics are also distances we take the inverse), and a distance metric based on the language family in Glottolog. The last metric denotes the number of steps one has to take to go from 1 language to another language in the language family tree.[9]

The scatterplot of the distance values against the error counts (Figure 6) show that larger distances for the metrics generally result in higher performance. The Pearson correlations confirm this with a correlation of -.11 for both lang2vec and language families (p=0.00), but only 0.01 for scripts (p=0.02), which is mostly due to a large number of high overlaps (Latin pairs). Although there seems to be a trend, the correlations are not very high, and accurate predictions of how many errors to expect

[7] https://en.wikipedia.org/wiki/List_of_Wikipedias

[8] Syntax and phonology categories are used, including KNN completions

[9] If languages are in different trees, we use $2 * max(depth\_lang1, depth\_lang2) + min(depth\_lang1, depth\_lang2)$ to ensure that out-of-family languages get a larger distance.

will be non-trivial based on these variables.

A manual inspection of the most common errors revealed some more interesting trends. We can clearly identify some challenging clusters of languages, where languages are closely related and the inputs might simply not contain enough information. This was the case for example for Dutch and its dialects, a South-Slavic cluster of Bosnian, Croatian, Serbian, and Montenegrin, and the Quechuan language family, especially within the Chinchay branch. Finally, we identified some errors in the data. Yakkha (Devanagari script) is very commonly overpredicted, even for utterances from other scripts (Latin), as it contains programming code in its training data. We also found that the language codes for Dai and Daai are probably confused, as they are a very common confusion for the model, but are not close in the glottolog tree. Finally, we found that TZM (Central Atlas Tamazight) and ZGH (Moroccan Amazigh) are they are commonly used interchangeably (they refer to the same language in Glottolog).

6 Conclusion

In this paper, we have compared simple n-gram frequency profiles with more complex competitors in a systematic variety of setups. We can confirm that the main finding of Cavnar et al. (1994) still holds: "Using N-gram frequency profiles provides a simple and reliable way to categorize documents in a wide range of classification tasks". Although more complex models outperform the n-gram frequency based model in in-domain setups, if we evaluate cross-domain character n-grams based on frequency heuristic or Naive Bayes have shown remarkable robustness. Cross-domain has also shown to be the most challenging setup, and is probably the most realistic setting, especially when including many languages for which we mainly have religious texts. We show that Multi-domain training partially resolves the performance drop. We also showed that using 1,000 utterances per language with a maximum of 100 characters per utterance already provides very good results. This finding has beneficial implications for future work; training can be done more efficiently, and smoothing over language labels might not be necessary, while more languages can be included. We also found that the number of languages and the amount of different language families or scripts are all not very influential to the performance.

Looking forward, we recommend to use language classifiers with care when they are applied out-of-domain, and where possible include a quantitative and qualitative analysis of its output (regardless of reported in-domain scores). Furthermore, we identify cross-domain language identification as the most prominent direction for future improvements.

Acknowledgements

I would like to express my gratitude to Ralf Brown for discussing and resolving labeling errors in the MIL-TALE data. Furthermore, I'd like to thank the NLPnorth members (and Mike of course). I also gladly thank the anonymous reviewers, who had some very interesting suggestions. Finally, thanks to Victor for showing me that Naive Bayes actually performs well for this task.

Limitations

We focus merely on language identification on the sentence/paragraph level, although there is also a stream of work on language identification on the word level which comes with its own set of challenges (Burchell et al., 2024), and has less annotated data available. Another constraint of our setup is that we assume one label per text. In some cases texts might be ambiguous with respect to language labels, i.e. they do not contain enough signal to discriminate between labels. Recently, there has been work in this direction with smaller sets of languages (Chifu et al., 2024; Fedorova et al., 2025).

We use language families as defined in Glottolog, only investigate scripts included in ISO-15924 and use language codes as defined in the ISO-639-3 standard, which is known to have issues and biases (Morey et al., 2013). We have used lang2vec for measuring typological distance, but it has shown to not be the most reliable solution for this task (Toossi et al., 2024).

We believe there are still a certain amount of incorrect labels in the data. Common issues are sentences from other languages being included (mainly English), wrong labels on the data-source level, and nonsensical sequences of characters. We believe that we have found most consistent errors in the labels through an automated detection of special script usage, followed by a manual inspection. However, our performance on English for example (85.20 F1) suggests that there are still some errors. Manual inspection revealed that overprediction on

English occured mainly on short sentences, sentences with names (of named entities), and quotes. Whereas utterances with other gold labels classified as English, are mainly from datasets where these phenomena occur in the training data.

References

Noëmi Aepli, Çağrı Çöltekin, Rob Van Der Goot, Tommi Jauhiainen, Mourhaf Kazzaz, Nikola Ljubešić, Kai North, Barbara Plank, Yves Scherrer, and Marcos Zampieri. 2023. Findings of the VarDial evaluation campaign 2023. In *Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023)*, pages 251–261, Dubrovnik, Croatia. Association for Computational Linguistics.

Milind Agarwal, Joshua Otten, and Antonios Anastasopoulos. 2025. Script-agnosticism and its impact on language identification for Dravidian languages. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 7364–7384, Albuquerque, New Mexico. Association for Computational Linguistics.

Rupal Bhargava, Yashvardhan Sharma, Shubham Sharma, and Abhinav Baid. 2015. Query labelling for indic languages using a hybrid approach. In *FIRE Workshops*, pages 40–42.

Ralf Brown. 2014. Non-linear mapping for improved identification of 1300+ languages. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 627–632, Doha, Qatar. Association for Computational Linguistics.

Laurie Burchell, Alexandra Birch, Nikolay Bogoychev, and Kenneth Heafield. 2023. An open dataset and model for language identification. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 865–879, Toronto, Canada. Association for Computational Linguistics.

Laurie Burchell, Alexandra Birch, Robert P Thompson, and Kenneth Heafield. 2024. Code-switched language identification is harder than you think. *arXiv preprint arXiv:2402.01505*.

Giacomo Camposampiero, Quynh Anh Nguyen, and Francesco Di Stefano. 2022. The curious case of logistic regression for Italian languages and dialects identification. In *Proceedings of the Ninth Workshop on NLP for Similar Languages, Varieties and Dialects*, pages 86–98, Gyeongju, Republic of Korea. Association for Computational Linguistics.

William B Cavnar, John M Trenkle, et al. 1994. N-gram-based text categorization. In *Proceedings of SDAIR-94, 3rd annual symposium on document analysis and information retrieval*, page 14, Las Vegas, NV.

Jordan Cazamias, Chinmayi Dixit, and Martina Marek. 2015. Large-scale language classification-writing a detector for 200 languages on twitter.

Wei-Rui Chen, Ife Adebara, Khai Duy Doan, Qisheng Liao, and Muhammad Abdul-Mageed. 2023. Fumbling in babel: An investigation into chatgpt's language identification ability. *arXiv preprint arXiv:2311.09696*.

Adrian-Gabriel Chifu, Goran Glavaš, Radu Tudor Ionescu, Nikola Ljubešić, Aleksandra Miletić, Filip Miletić, Yves Scherrer, and Ivan Vulić. 2024. VarDial evaluation campaign 2024: Commonsense reasoning in dialects and multi-label similar language identification. In *Proceedings of the Eleventh Workshop on NLP for Similar Languages, Varieties, and Dialects (VarDial 2024)*, pages 1–15, Mexico City, Mexico. Association for Computational Linguistics.

Alexis Conneau and Guillaume Lample. 2019. Cross-lingual language model pretraining. *Advances in neural information processing systems*, 32.

A. Seza Doğruöz, Sunayana Sitaram, Barbara E. Bullock, and Almeida Jacqueline Toribio. 2021. A survey of code-switching: Linguistic and social perspectives for language technologies. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1654–1666, Online. Association for Computational Linguistics.

Mariia Fedorova, Jonas Sebulon Frydenberg, Victoria Handford, Victoria Ovedie Chruickshank Langø, Solveig Helene Willoch, Marthe Løken Midtgaard, Yves Scherrer, Petter Mæhlum, and David Samuel. 2025. Multi-label Scandinavian language identification (SLIDE). In *Proceedings of the Third Workshop on Resources and Representations for Under-Resourced Languages and Domains (RESOURCEFUL-2025)*, pages 179–189, Tallinn, Estonia. University of Tartu Library, Estonia.

Alex Graves and Jürgen Schmidhuber. 2005. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610.

Harald Hammarström, Robert Forkel, Martin Haspelmath, and Sebastian Bank. 2023. Glottolog 4.8. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Ayyoob ImaniGooghari, 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.

Tommi Jauhiainen, Marco Lui, Marcos Zampieri, Timothy Baldwin, and Krister Lindén. 2019. Automatic language identification in texts: A survey. *Journal of Artificial Intelligence Research*, 65:675–782.

Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 427–431, Valencia, Spain. Association for Computational Linguistics.

Amir Kargaran, Ayyoob Imani, François Yvon, and Hinrich Schuetze. 2023a. GlotLID: Language identification for low-resource languages. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 6155–6218, Singapore. Association for Computational Linguistics.

Amir Hossein Kargaran, François Yvon, and Hinrich Schütze. 2023b. Glotscript: A resource and tool for low resource writing system identification. *arXiv preprint arXiv:2309.13320*.

Julia Kreutzer, Isaac Caswell, Lisa Wang, Ahsan Wahab, Daan van Esch, Nasanbayar Ulzii-Orshikh, Allahsera Tapo, Nishant Subramani, Artem Sokolov, Claytone Sikasote, Monang Setyawan, Supheakmungkol Sarin, Sokhar Samb, Benoît Sagot, Clara Rivera, Annette Rios, Isabel Papadimitriou, Salomey Osei, Pedro Ortiz Suarez, Iroro Orife, Kelechi Ogueji, Andre Niyongabo Rubungo, Toan Q. Nguyen, Mathias Müller, André Müller, Shamsuddeen Hassan Muhammad, Nanda Muhammad, Ayanda Mnyakeni, Jamshidbek Mirzakhalov, Tapiwanashe Matangira, Colin Leong, Nze Lawson, Sneha Kudugunta, Yacine Jernite, Mathias Jenny, Orhan Firat, Bonaventure F. P. Dossou, Sakhile Dlamini, Nisansa de Silva, Sakine Çabuk Ballı, Stella Biderman, Alessia Battisti, Ahmed Baruwa, Ankur Bapna, Pallavi Baljekar, Israel Abebe Azime, Ayodele Awokoya, Duygu Ataman, Orevaoghene Ahia, Oghenefego Ahia, Sweta Agrawal, and Mofetoluwa Adeyemi. 2022. Quality at a glance: An audit of web-crawled multilingual datasets. *Transactions of the Association for Computational Linguistics*, 10:50–72.

Patrick Littell, David R. Mortensen, Ke Lin, Katherine Kairis, Carlisle Turner, and Lori Levin. 2017. URIEL and lang2vec: Representing languages as typological, geographical, and phylogenetic vectors. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 8–14, Valencia, Spain. Association for Computational Linguistics.

Marco Lui and Timothy Baldwin. 2011. Cross-domain feature selection for language identification. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 553–561, Chiang Mai, Thailand. Asian Federation of Natural Language Processing.

Marco Lui and Timothy Baldwin. 2014a. Accurate language identification of Twitter messages. In *Proceedings of the 5th Workshop on Language Analysis for Social Media (LASM)*, pages 17–25, Gothenburg, Sweden. Association for Computational Linguistics.

Marco Lui and Timothy Baldwin. 2014b. Accurate language identification of twitter messages. In *Proceedings of the 5th Workshop on Language Analysis for Social Media (LASM)*, pages 17–25, Gothenburg, Sweden. Association for Computational Linguistics.

Martin Majliš. 2012. Yet another language identifier. In *Proceedings of the Student Research Workshop at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 46–54, Avignon, France. Association for Computational Linguistics.

Stephen Morey, Mark W Post, and Victor A Friedman. 2013. The language codes of iso 639: A premature, ultimately unobtainable, and possibly damaging standardization. *PARADISEC RRR Conference*.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Nathaniel Robinson, Perez Ogayo, David R. Mortensen, and Graham Neubig. 2023. ChatGPT MT: Competitive for high- (but not low-) resource languages. In *Proceedings of the Eighth Conference on Machine Translation*, pages 392–418, Singapore. Association for Computational Linguistics.

NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, Anna Sun, Skyler Wang, Guillaume Wenzek, Al Youngblood, Bapi Akula, Loic Barrault, Gabriel Mejia Gonzalez, Prangthip Hansanti, John Hoffman, Semarley Jarrett, Kaushik Ram Sadagopan, Dirk Rowe, Shannon Spruit, Chau Tran, Pierre Andrews, Necip Fazil Ayan, Shruti Bhosale, Sergey Edunov, Angela Fan, Cynthia Gao, Vedanuj Goswami, Francisco Guzmán, Philipp Koehn, Alexandre Mourachko, Christophe Ropers, Safiyyah Saleem, Holger Schwenk, and Jeff Wang. 2022. No language left behind: Scaling human-centered machine translation. *arXiv preprint arXiv:2207.04672*.

Mads Toftrup, Søren Asger Sørensen, Manuel R. Ciosici, and Ira Assent. 2021. A reproduction of apple's bi-directional LSTM models for language identification in short strings. In *Proceedings of the*

16th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop, pages 36–42, Online. Association for Computational Linguistics.

Hasti Toossi, Guo Huai, Jinyu Liu, Eric Khiu, A. Seza Doğruöz, and En-Shiun Lee. 2024. A reproducibility study on quantifying language similarity: The impact of missing values in the URIEL knowledge base. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop)*, pages 233–241, Mexico City, Mexico. Association for Computational Linguistics.

Gertjan van Noord. 1997. Textcat.

Daniel Varab and Natalie Schluter. 2021. MassiveSumm: a very large-scale, very multilingual, news summarisation dataset. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10150–10161, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Genta Winata, Alham Fikri Aji, Zheng Xin Yong, and Thamar Solorio. 2023. The decades progress on code-switching research in NLP: A systematic survey on trends and challenges. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 2936–2978, Toronto, Canada. Association for Computational Linguistics.

Marcos Zampieri, Liling Tan, Nikola Ljubešić, and Jörg Tiedemann. 2014. A report on the DSL shared task 2014. In *Proceedings of the First Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects*, pages 58–67, Dublin, Ireland. Association for Computational Linguistics and Dublin City University.

Marcos Zampieri, Liling Tan, Nikola Ljubešić, Jörg Tiedemann, and Preslav Nakov. 2015. Overview of the DSL shared task 2015. In *Proceedings of the Joint Workshop on Language Technology for Closely Related Languages, Varieties and Dialects*, pages 1–9, Hissar, Bulgaria. Association for Computational Linguistics.

Daniel Zeman, Joakim Nivre, Mitchell Abrams, Elia Ackermann, Noëmi Aepli, Hamid Aghaei, Željko Agić, Amir Ahmadi, Lars Ahrenberg, Chika Kennedy Ajede, Salih Furkan Akkurt, Gabrielė Aleksandravičiūtė, Ika Alfina, Avner Algom, Khalid Alnajjar, Chiara Alzetta, Erik Andersen, Lene Antonsen, Tatsuya Aoyama, Katya Aplonova, Angelina Aquino, Carolina Aragon, Glyd Aranes, Maria Jesus Aranzabe, Bilge Nas Arıcan, H͡órunn Arnardóttir, Gashaw Arutie, Jessica Naraiswari Arwidarasti, Masayuki Asahara, Katla Ásgeirsdóttir, Deniz Baran Aslan, Cengiz Asmazoğlu, Luma Ateyah, Furkan Atmaca, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Mariana Avelãs, Elena Badmaeva, Keerthana Balasubramani, Miguel Ballesteros, Esha Banerjee, Sebastian Bank, Verginica Barbu Mititelu, Starkaður Barkarson, Rodolfo Basile, Victoria Basmov, Colin Batchelor, John Bauer, Seyyit Talha Bedir, Shabnam Behzad, Kepa Bengoetxea, İbrahim Benli, Yifat Ben Moshe, Gözde Berk, Riyaz Ahmad Bhat, Erica Biagetti, Eckhard Bick, Agnė Bielinskienė, Kristín Bjarnadóttir, Rogier Blokland, Victoria Bobicev, Loïc Boizou, Emanuel Borges Völker, Carl Börstell, Cristina Bosco, Gosse Bouma, Sam Bowman, Adriane Boyd, Anouck Braggaar, António Branco, Kristina Brokaitė, Aljoscha Burchardt, Marisa Campos, Marie Candito, Bernard Caron, Gauthier Caron, Catarina Carvalheiro, Rita Carvalho, Lauren Cassidy, Maria Clara Castro, Sérgio Castro, Tatiana Cavalcanti, Gülşen Cebiroğlu Eryiğit, Flavio Massimiliano Cecchini, Giuseppe G. A. Celano, Slavomír Čéplö, Neslihan Cesur, Savas Cetin, Özlem Çetinoğlu, Fabricio Chalub, Liyanage Chamila, Shweta Chauhan, Ethan Chi, Taishi Chika, Yongseok Cho, Jinho Choi, Jayeol Chun, Juyeon Chung, Alessandra T. Cignarella, Silvie Cinková, Aurélie Collomb, Çağrı Çöltekin, Miriam Connor, Daniela Corbetta, Francisco Costa, Marine Courtin, Mihaela Cristescu, Ingerid Løyning Dale, Philemon Daniel, Elizabeth Davidson, Leonel Figueiredo de Alencar, Mathieu Dehouck, Martina de Laurentiis, Marie-Catherine de Marneffe, Valeria de Paiva, Mehmet Oguz Derin, Elvis de Souza, Arantza Diaz de Ilarraza, Carly Dickerson, Arawinda Dinakaramani, Elisa Di Nuovo, Bamba Dione, Peter Dirix, Kaja Dobrovoljc, Adrian Doyle, Timothy Dozat, Kira Droganova, Puneet Dwivedi, Christian Ebert, Hanne Eckhoff, Masaki Eguchi, Sandra Eiche, Marhaba Eli, Ali Elkahky, Binyam Ephrem, Olga Erina, Tomaž Erjavec, Farah Essaidi, Aline Etienne, Wograine Evelyn, Sidney Facundes, Richárd Farkas, Federica Favero, Jannatul Ferdaousi, Marília Fernanda, Hector Fernandez Alcalde, Amal Fethi, Jennifer Foster, Cláudia Freitas, Kazunori Fujita, Katarína Gajdošová, Daniel Galbraith, Federica Gamba, Marcos Garcia, Moa Gärdenfors, Fabrício Ferraz Gerardi, Kim Gerdes, Luke Gessler, Filip Ginter, Gustavo Godoy, Iakes Goenaga, Koldo Gojenola, Memduh Gökırmak, Yoav Goldberg, Xavier Gómez Guinovart, Berta González Saavedra, Bernadeta Griciūtė, Matias Grioni, Loïc Grobol, Normunds Grūzītis, Bruno Guillaume, Céline Guillot-Barbance, Tunga Güngör, Nizar Habash, Hinrik Hafsteinsson, Jan Hajič, Jan Hajič jr., Mika Hämäläinen, Linh Hà Mỹ, Na-Rae Han, Muhammad Yudistira Hanifmuti, Takahiro Harada, Sam Hardwick, Kim Harris, Dag Haug, Johannes Heinecke, Oliver Hellwig, Felix Hennig, Barbora Hladká, Jaroslava Hlaváčová, Florinel Hociung, Petter Hohle, Marivel Huerta Mendez, Jena Hwang, Takumi Ikeda, Anton Karl Ingason, Radu Ion, Elena Irimia, Ọlájídé Ishola, Artan Islamaj, Kaoru Ito, Siratun Jannat, Tomáš Jelínek, Apoorva Jha, Katharine Jiang, Anders Johannsen, Hildur Jónsdóttir, Fredrik Jørgensen, Markus Juutinen, Hüner Kaşıkara, Nadezhda Kabaeva, Sylvain Kahane, Hiroshi Kanayama, Jenna Kanerva, Neslihan Kara, Ritván Karahóǧa, Andre Kåsen, Tolga Kayadelen, Sarveswaran Kengatharaiyer, Václava Kettnerová, Jesse Kirchner, Elena

Klementieva, Elena Klyachko, Arne Köhn, Abdullatif Köksal, Kamil Kopacewicz, Timo Korkiakangas, Mehmet Köse, Alexey Koshevoy, Natalia Kotsyba, Jolanta Kovalevskaitė, Simon Krek, Parameswari Krishnamurthy, Sandra Kübler, Adrian Kuqi, Oğuzhan Kuyrukçu, Aslı Kuzgun, Sookyoung Kwak, Kris Kyle, Veronika Laippala, Lorenzo Lambertino, Tatiana Lando, Septina Dian Larasati, Alexei Lavrentiev, John Lee, Phuong Lê Hồng, Alessandro Lenci, Saran Lertpradit, Herman Leung, Maria Levina, Lauren Levine, Cheuk Ying Li, Josie Li, Keying Li, Yixuan Li, Yuan Li, KyungTae Lim, Bruna Lima Padovani, Yi-Ju Jessica Lin, Krister Lindén, Yang Janet Liu, Nikola Ljubešić, Olga Loginova, Stefano Lusito, Andry Luthfi, Mikko Luukko, Olga Lyashevskaya, Teresa Lynn, Vivien Macketanz, Menel Mahamdi, Jean Maillard, Ilya Makarchuk, Aibek Makazhanov, Michael Mandl, Christopher Manning, Ruli Manurung, Büşra Marşan, Cătălina Mărănduc, David Mareček, Katrin Marheinecke, Stella Markantonatou, Héctor Martínez Alonso, Lorena Martín Rodríguez, André Martins, Cláudia Martins, Jan Mašek, Hiroshi Matsuda, Yuji Matsumoto, Alessandro Mazzei, Ryan McDonald, Sarah McGuinness, Gustavo Mendonça, Tatiana Merzhevich, Niko Miekka, Aaron Miller, Karina Mischenkova, Anna Missilä, Cătălin Mititelu, Maria Mitrofan, Yusuke Miyao, AmirHossein Mojiri Foroushani, Judit Molnár, Amirsaeid Moloodi, Simonetta Montemagni, Amir More, Laura Moreno Romero, Giovanni Moretti, Shinsuke Mori, Tomohiko Morioka, Shigeki Moro, Bjartur Mortensen, Bohdan Moskalevskyi, Kadri Muischnek, Robert Munro, Yugo Murawaki, Kaili Müürisep, Pinkey Nainwani, Mariam Nakhlé, Juan Ignacio Navarro Horñiacek, Anna Nedoluzhko, Gunta Nešpore-Bērzkalne, Manuela Nevaci, Luong Nguyễn Thị, Huyền Nguyễn Thị Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Alireza Nourian, Hanna Nurmi, Stina Ojala, Atul Kr. Ojha, Hulda Óladóttir, Adédayọ Olúòkun, Mai Omura, Emeka Onwuegbuzia, Noam Ordan, Petya Osenova, Robert Östling, Lilja Øvrelid, Şaziye Betül Özateş, Merve Özçelik, Arzucan Özgür, Balkız Öztürk Başaran, Teresa Paccosi, Alessio Palmero Aprosio, Anastasia Panova, Hyunji Hayley Park, Niko Partanen, Elena Pascual, Marco Passarotti, Agnieszka Patejuk, Guilherme Paulino-Passos, Giulia Pedonese, Angelika Peljak-Łapińska, Siyao Peng, Siyao Logan Peng, Rita Pereira, Sílvia Pereira, Cenel-Augusto Perez, Natalia Perkova, Guy Perrier, Slav Petrov, Daria Petrova, Andrea Peverelli, Jason Phelan, Jussi Piitulainen, Yuval Pinter, Clara Pinto, Tommi A Pirinen, Emily Pitler, Magdalena Plamada, Barbara Plank, Thierry Poibeau, Larisa Ponomareva, Martin Popel, Lauma Pretkalniņa, Sophie Prévost, Prokopis Prokopidis, Adam Przepiórkowski, Robert Pugh, Tiina Puolakainen, Sampo Pyysalo, Peng Qi, Andreia Querido, Andriela Rääbis, Alexandre Rademaker, Mizanur Rahoman, Taraka Rama, Loganathan Ramasamy, Joana Ramos, Fam Rashel, Mohammad Sadegh Rasooli, Vinit Ravishankar, Livy Real, Petru Rebeja, Siva Reddy, Mathilde Regnault, Georg Rehm, Arij Riabi, Ivan Riabov, Michael Rießler, Erika Rimkutė, Larissa Rinaldi, Laura Rituma, Putri Rizqiyah, Luisa Rocha, Eiríkur Rögnvaldsson, Ivan Roksandic, Mykhailo Romanenko, Rudolf Rosa, Valentin Roșca, Davide Rovati, Ben Rozonoyer, Olga Rudina, Jack Rueter, Kristján Rúnarsson, Shoval Sadde, Pegah Safari, Aleksi Sahala, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Stephanie Samson, Manuela Sanguinetti, Ezgi Sanıyar, Dage Särg, Marta Sartor, Mitsuya Sasaki, Baiba Saulīte, Yanin Sawanakunanon, Shefali Saxena, Kevin Scannell, Salvatore Scarlata, Nathan Schneider, Sebastian Schuster, Lane Schwartz, Djamé Seddah, Wolfgang Seeker, Mojgan Seraji, Syeda Shahzadi, Mo Shen, Atsuko Shimada, Hiroyuki Shirasu, Yana Shishkina, Muh Shohibussirri, Maria Shvedova, Janine Siewert, Einar Freyr Sigurðsson, João Silva, Aline Silveira, Natalia Silveira, Sara Silveira, Maria Simi, Radu Simionescu, Katalin Simkó, Mária Šimková, Haukur Barri Símonarson, Kiril Simov, Dmitri Sitchinava, Ted Sither, Maria Skachedubova, Aaron Smith, Isabela Soares-Bastos, Per Erik Solberg, Barbara Sonnenhauser, Shafi Sourov, Rachele Sprugnoli, Vivian Stamou, Steinh͡ór Steingrímsson, Antonio Stella, Abishek Stephen, Milan Straka, Emmett Strickland, Jana Strnadová, Alane Suhr, Yogi Lesmana Sulestio, Umut Sulubacak, Shingo Suzuki, Daniel Swanson, Zsolt Szántó, Chihiro Taguchi, Dima Taji, Fabio Tamburini, Mary Ann C. Tan, Takaaki Tanaka, Dipta Tanaya, Mirko Tavoni, Samson Tella, Isabelle Tellier, Marinella Testori, Guillaume Thomas, Sara Tonelli, Liisi Torga, Marsida Toska, Trond Trosterud, Anna Trukhina, Reut Tsarfaty, Utku Türk, Francis Tyers, Sveinbjörn H͡órðarson, Vilhjálmur H͡orsteinsson, Sumire Uematsu, Roman Untilov, Zdeňka Urešová, Larraitz Uria, Hans Uszkoreit, Andrius Utka, Elena Vagnoni, Sowmya Vajjala, Socrates Vak, Rob van der Goot, Martine Vanhove, Daniel van Niekerk, Gertjan van Noord, Viktor Varga, Uliana Vedenina, Giulia Venturi, Veronika Vincze, Natalia Vlasova, Aya Wakasa, Joel C. Wallenberg, Lars Wallin, Abigail Walsh, Jonathan North Washington, Maximilan Wendt, Paul Widmer, Shira Wigderson, Sri Hartati Wijono, Seyi Williams, Mats Wirén, Christian Wittern, Tsegay Woldemariam, Tak-sum Wong, Alina Wróblewska, Mary Yako, Kayo Yamashita, Naoki Yamazaki, Chunxiao Yan, Koichi Yasuoka, Marat M. Yavrumyan, Arife Betül Yenice, Olcay Taner Yıldız, Zhuoran Yu, Arlisa Yuliawati, Zdeněk Žabokrtský, Shorouq Zahra, Amir Zeldes, He Zhou, Hanzhi Zhu, Yilun Zhu, Anna Zhuravleva, and Rayan Ziane. 2023. Universal dependencies 2.12. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.

parameter	range
LR	**0.0001**, 0.00001
Batch size	**16**, 32, 64
Dropout	**0.0**, 0.2, 0.3
Hidden size	120, **768**
Num layers	1, **2**

Table 4: Hyperparameter ranges evaluated for the LSTM model, best ones are in bold, and are the ones used in the paper.

A LSTM tuning

We tuned the most important hyperparameters for our LSTM model on a set of 100 random languages. The ranges we evaluated are reported in Table 4.

B Standard deviation for size experiments

In Figure 7 we report the standard deviation for all experiments of Section 4.1.

C Language models comparison

We compared the performance of a variety of multilingual language models on an earlier release of the MIL-TALE dataset (including 1,277 languages), without any constraints like character size limitations. Figure 8 shows that GLOT500 performs slightly favorable compared to its competitors. The Byt5 model crashed during training, but performance was substantially lower and these experiments are computationally costly, so we never finished the full training procedure.

D Cleaning procedure

1. We group all utterances of each language label of each dataset; we clean only on the label level (i.e. we only remove a language completely), and not on the instance level to keep the data more realistic.

2. We remove all data from dialects, macro-languages, and datasets with language codes not found in the ISO-639-3 standard. Expired language codes are mapped, following the official 639-3 updates.

3. We standardize the domain labels by creating a mapping.

4. We find the most common script based on the ISO 15924 standard (same as Unicode), and check whether this script is supposed to be used for this language with the data from Kargaran et al. (2023b). Note that this also catches some erroneous encoding issues etc. (as the script would then be 'Common' for example.

5. We manually inspect the texts of odd combinations and frequencies of scripts, and found some XML labels and other markup. We either corrected these in the original data source, or in our pre-processing scripts.

E Per language performance

We report the F1 scores and precision-recall for each language in Table 5 until Table 10. These results are from the 2,034 settings, where we include all languages as a label (so 1 label can have multiple scripts).

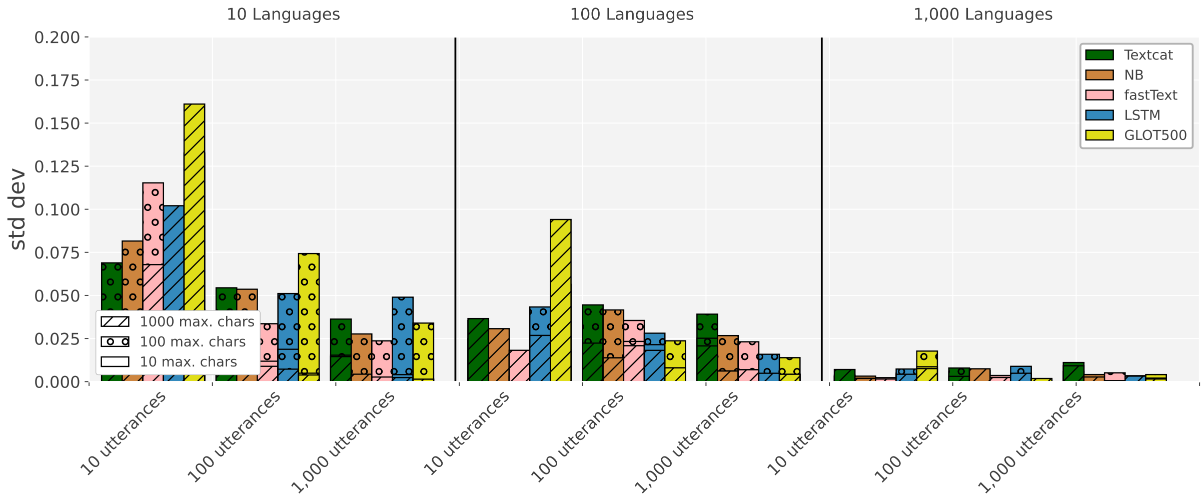


Figure 7: Standard deviations for all experiments with different sizes (equivalent structure as Figure 2).

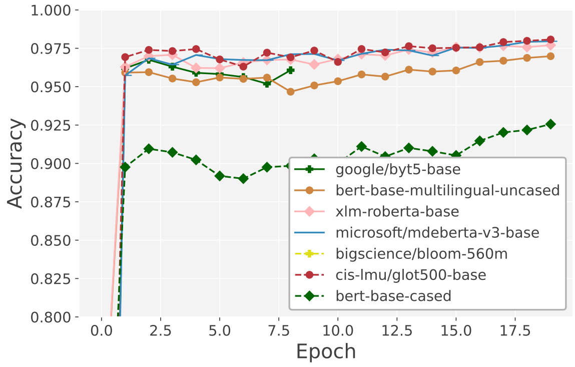


Figure 8: Comparison of different language models over time (i.e. number of epochs).

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
aaa	98.5	-2.9	84.4	-27.0	99.7	-0.5	99.6	-0.5	99.7	0.0
aah	96.8	-3.3	99.9	-0.3	99.7	-0.4	96.6	0.8	99.5	-0.8
aai	95.8	2.1	98.4	-0.4	96.6	3.0	94.8	0.9	97.9	1.7
aak	99.9	-0.3	99.0	-1.9	100.0	-0.0	100.0	0.0	100.0	0.0
aar	69.8	20.4	90.2	-7.9	84.5	10.3	78.1	16.3	87.5	11.5
aau	98.7	-1.3	99.6	-0.9	99.6	-0.2	99.5	0.6	99.5	-0.3
aaz	87.8	15.4	96.3	-0.1	92.1	10.9	90.9	14.3	94.1	8.9
abi	93.5	11.9	96.9	5.0	86.2	24.0	70.8	45.0	85.3	25.5
abk	81.2	25.4	72.7	39.2	85.8	13.1	85.6	16.1	89.3	15.6
abn	94.8	3.4	94.0	9.6	96.2	4.1	96.3	0.6	98.1	1.0
abs	88.7	2.4	93.9	7.4	94.7	-5.3	91.2	-10.8	94.8	-7.0
abx	94.3	0.2	97.9	-0.9	96.8	3.2	96.1	5.7	97.7	2.3
aby	96.4	-5.1	98.8	-2.3	98.8	-1.1	97.5	-0.5	99.0	-0.5
aca	99.1	-1.1	98.4	-3.1	98.9	-1.8	99.3	-0.5	99.9	-0.3
ace	85.4	16.9	88.2	12.5	84.7	6.5	83.6	9.9	89.1	11.0
acf	97.6	0.5	99.4	-0.4	97.6	1.7	96.8	1.5	98.7	1.1
ach	85.4	10.5	91.9	10.0	91.4	4.3	85.3	9.3	92.2	3.7
acm	8.1	11.2	0.0	0.0	35.2	32.8	21.0	16.5	37.5	11.5
acn	90.9	12.6	92.5	13.2	90.7	7.6	86.9	-1.4	93.1	5.7
acr	95.1	4.4	99.0	0.4	97.6	-0.3	97.3	0.1	98.8	-0.4
acu	87.0	-4.0	96.1	-2.7	94.4	0.1	87.1	4.5	94.2	-0.3
ada	88.7	10.1	92.8	12.1	92.5	7.6	90.2	4.9	94.9	6.1
ade	84.9	18.5	86.9	21.8	84.8	20.5	79.4	5.4	85.5	18.9
adh	88.6	10.5	93.9	9.2	89.9	1.0	83.8	-4.8	91.3	-4.5
adi	91.7	8.3	93.8	4.0	90.9	2.6	91.3	6.3	93.6	4.1
adj	96.1	-6.4	97.4	-4.9	98.5	-2.4	99.7	0.0	99.3	-0.9
adl	85.9	18.7	83.4	3.2	80.0	-13.0	81.1	-6.2	82.2	-10.5
adx	41.6	6.4	51.2	-8.9	47.4	15.0	36.5	28.7	57.6	-19.4
ady	81.5	4.5	84.9	8.3	86.7	-11.2	77.5	-24.7	88.6	-11.0
aeb	33.6	-0.4	11.0	59.0	69.4	-17.6	31.6	3.7	74.4	-5.2
aer	99.2	-1.6	98.6	-2.8	99.7	-0.4	99.9	0.1	99.9	0.0
aeu	99.4	-0.5	99.2	-1.7	99.6	-0.1	99.5	-0.7	99.5	-0.6
aey	96.1	-4.1	99.1	-0.6	97.4	-0.2	97.8	0.8	98.7	1.1
afr	84.5	-1.1	88.1	-0.1	92.2	-7.5	84.8	-10.3	96.8	0.5
agd	97.4	-4.7	99.4	-0.8	99.0	0.4	99.0	1.0	99.2	-0.1
agg	98.2	2.3	97.8	1.3	98.5	0.2	98.7	2.0	99.0	1.7
agm	99.2	-1.7	99.2	-1.7	99.7	-0.5	99.9	-0.2	99.9	-0.1
agn	92.8	-10.1	98.5	-1.6	98.4	-1.4	97.9	0.3	98.8	-0.4
agq	89.1	17.9	90.3	12.6	93.5	2.8	92.6	3.1	93.4	0.0
agr	86.5	3.7	96.2	3.4	92.1	6.9	85.7	13.6	91.4	7.2
ags	97.6	3.4	98.3	2.7	96.2	-0.5	96.9	1.1	96.6	-0.7
agt	93.3	-4.6	98.5	-1.1	97.7	0.6	96.8	2.8	98.2	2.0
agu	96.6	-1.2	99.6	-0.2	96.1	4.5	93.0	9.7	98.0	1.9
agw	94.7	-3.4	98.3	2.3	96.8	-1.3	96.0	-4.6	98.2	-1.6
aha	97.3	-4.7	98.7	-2.6	96.6	-6.1	97.1	-5.1	96.5	-6.7
ahk	99.3	0.2	98.3	1.4	99.7	0.2	99.8	-0.0	99.8	-0.5
aia	97.5	-0.7	98.7	-0.5	98.3	-0.4	97.5	-0.1	98.7	0.2
aii	99.3	-1.4	99.1	-1.8	99.9	-0.2	99.9	-0.0	99.3	-1.5
aim	79.5	-1.9	89.4	-7.9	88.0	-8.5	80.5	-19.5	83.2	-20.0
ain	94.6	-2.6	97.9	1.9	97.0	1.1	94.5	-2.8	97.2	1.5
aji	98.1	3.2	97.9	1.7	95.9	-0.9	97.5	0.8	97.9	0.3
ajz	89.5	4.7	96.9	2.4	93.8	-3.2	92.1	-1.7	93.1	-4.1
akb	58.7	13.2	77.2	24.5	71.5	25.7	62.1	25.9	67.4	-15.9
ake	98.4	2.8	99.3	0.7	98.5	1.9	98.3	2.5	98.7	1.6
akh	98.9	-1.9	99.7	-0.6	99.9	-0.1	99.9	0.2	99.9	-0.1
ald	98.1	0.3	98.5	-0.4	98.7	1.1	98.1	0.3	99.0	0.8
alj	97.2	-0.1	95.4	-8.8	99.1	-0.3	98.9	-0.0	99.6	-0.3
alp	91.8	-4.4	98.1	-0.7	96.3	1.0	93.8	6.0	97.1	2.0
alq	95.7	4.2	95.4	1.1	94.6	-0.4	94.7	-0.8	97.0	1.8
als	93.0	-9.6	93.4	-11.1	98.4	-0.6	98.7	1.4	99.0	0.7
alt	84.0	-8.9	92.9	1.8	95.5	-1.0	91.0	-9.6	96.8	-1.6
alw	85.8	-4.7	86.9	-16.4	98.8	1.0	94.3	-0.7	98.3	0.1
aly	96.0	0.0	95.7	-3.9	98.7	1.7	96.8	4.6	98.2	0.5
alz	89.6	7.0	92.3	7.4	93.6	5.6	90.7	-0.5	94.6	4.0
ame	98.9	-0.5	97.6	-3.3	99.1	0.9	99.5	0.8	99.6	0.7
amf	92.3	-4.5	98.4	0.8	96.8	2.5	93.2	-0.4	96.2	2.0
amh	93.6	-9.5	68.1	-47.7	98.0	-1.8	95.1	-7.1	99.3	-1.2
ami	78.2	4.2	87.6	18.6	83.8	5.2	81.9	3.5	86.0	3.6
amk	95.9	0.1	98.6	-1.2	97.2	-0.6	96.4	0.6	97.9	-0.7
amm	97.5	-4.9	97.7	-4.5	99.6	-0.6	99.8	0.0	99.6	-0.3
amo	91.1	-14.1	99.3	-1.0	98.3	-2.2	96.4	-1.0	98.2	-1.0
amp	98.0	-0.5	96.3	-5.3	98.3	1.1	98.2	1.7	98.9	1.4
amr	99.4	-1.1	99.2	-1.5	99.6	0.3	99.8	0.2	99.9	0.0
amu	98.3	3.0	99.4	0.1	99.2	0.3	98.9	1.4	99.2	0.5
amx	97.5	-2.8	97.7	-4.5	98.7	-1.5	96.6	-6.0	98.6	-1.7
ang	97.3	-2.1	94.5	-7.5	94.8	-4.7	98.6	1.2	97.1	-3.6
anm	94.8	8.1	96.5	5.9	95.1	3.9	92.8	-0.8	96.3	1.7
ann	96.3	2.1	97.8	1.3	97.1	1.9	97.5	0.8	98.1	1.1
anp	25.8	-16.7	52.4	-32.5	63.0	-8.5	44.2	-21.5	63.9	-36.4
anv	99.4	-0.9	100.0	-0.1	99.6	-0.3	99.0	-1.7	99.8	-0.4
aoi	99.7	-0.5	98.3	-3.3	99.4	-0.8	99.7	0.2	99.6	-0.3
aoj	99.3	-1.3	98.9	-1.7	99.4	-0.5	99.6	0.0	100.0	-0.1
aom	99.5	-0.2	98.7	-2.1	99.9	0.1	99.7	0.6	99.8	0.2
aon	96.9	-2.1	99.3	-1.1	98.1	-0.8	97.8	0.6	98.9	-0.3
aoz	90.2	-5.6	95.8	3.5	93.9	-0.7	92.7	0.7	95.0	-2.3
apb	97.0	0.3	98.9	1.0	98.1	0.8	97.2	2.2	98.4	1.9
apc	35.7	-1.2	14.4	48.1	64.5	14.5	17.1	65.2	67.2	23.0
apd	64.3	-32.5	59.0	-58.1	76.2	-28.1	73.1	-21.2	73.9	-24.5
ape	98.9	-2.1	99.4	-1.3	99.7	-0.7	99.5	-0.4	99.8	-0.4
apf	91.6	-6.7	99.0	-1.6	98.1	0.9	97.2	2.5	99.3	1.0
apn	99.0	-1.4	99.8	-0.5	99.5	0.4	99.5	-0.1	99.7	0.0
apr	96.2	-2.3	99.4	-0.0	98.5	1.7	97.3	2.2	99.0	0.9
aps	94.5	1.8	96.2	1.5	95.7	2.0	91.6	4.0	93.9	-3.7
apt	45.0	-51.1	48.3	-56.7	87.4	-0.2	88.8	5.5	90.9	0.6
apu	97.8	-3.4	87.0	-22.6	98.6	-0.8	98.9	0.7	98.8	-0.4
apw	99.4	-0.9	97.8	-4.3	99.8	-0.5	99.9	0.2	100.0	0.0
apy	96.1	-4.8	99.3	-1.3	98.6	0.2	98.3	0.7	99.0	1.3
apz	99.4	-1.2	98.7	-2.6	99.3	-0.9	99.7	-0.5	99.7	-0.4
arb	3.4	17.9	18.5	19.3	77.7	-2.4	43.4	-47.8	72.1	-25.2
arc	99.2	1.6	96.9	-1.3	99.4	0.8	99.6	0.7	99.1	1.8
are	97.9	2.4	99.0	0.3	96.8	5.9	94.2	10.2	96.1	7.0
arg	67.4	-9.5	80.9	-25.2	92.0	1.6	79.0	29.3	96.0	6.1
arl	99.4	-1.1	98.7	-2.5	99.6	-0.6	99.7	-0.2	99.0	-1.6
arn	98.9	1.2	98.6	-1.1	97.7	-1.9	98.1	-0.1	99.3	0.4
arq	48.5	-35.9	34.5	-71.9	78.8	-2.0	44.5	42.6	78.7	17.9
ars	33.0	-14.3	32.1	12.0	60.5	2.0	27.9	18.8	63.5	5.7
ary	59.4	0.1	64.1	31.5	81.9	-13.4	56.2	-49.6	84.2	-14.4
arz	33.2	-8.8	31.5	-0.1	61.0	3.8	24.8	38.7	65.3	10.7
asa	74.0	-9.0	91.9	4.5	86.5	-7.7	76.4	-13.6	85.8	-13.0
asc	73.3	27.1	76.1	33.2	74.0	10.0	70.7	15.0	72.5	5.0
asg	86.1	14.6	96.2	6.1	91.5	2.2	85.5	-1.0	90.8	-5.4
asm	92.3	6.5	97.5	1.9	98.7	-0.5	98.0	1.5	98.9	1.0
aso	99.1	0.2	99.0	-1.1	98.8	1.5	98.9	1.4	98.7	1.0
ast	73.2	-11.5	88.0	-14.0	89.3	-3.0	80.7	20.8	95.2	-0.1
ata	93.9	8.3	95.8	5.2	95.5	2.8	95.3	3.6	97.4	2.2
atb	98.5	0.8	99.6	0.5	98.9	0.8	99.2	1.4	99.6	0.7
atd	93.3	-2.5	99.3	-1.1	98.6	-1.8	96.5	-2.5	99.4	0.4
atg	94.1	3.0	98.0	2.5	96.1	2.4	95.2	2.2	96.8	1.3
ati	99.9	-0.2	99.6	-0.8	99.9	-0.1	100.0	0.0	99.9	0.1
atj	75.2	-10.7	82.7	-0.2	86.5	-4.8	85.1	-3.7	83.6	-15.6
atm	93.3	-9.0	95.7	-8.0	96.2	-2.9	95.8	-2.4	97.7	-1.4
atq	86.1	2.3	95.1	4.9	91.9	0.9	88.4	1.5	93.0	2.3
att	98.7	-0.5	99.5	-0.5	98.9	1.1	98.4	2.3	99.1	0.8
aty	71.4	-9.1	86.3	-16.5	86.8	-8.9	80.7	-11.2	83.8	-20.5
auc	99.8	0.0	99.6	-0.5	99.8	-0.3	99.7	0.0	99.6	0.1
aui	91.5	-9.5	98.8	-2.0	97.4	-1.8	95.0	-3.7	97.9	-0.9
auy	97.5	3.2	98.6	-0.1	97.9	1.0	97.4	0.3	98.4	0.7
ava	88.3	3.3	89.3	-6.2	93.8	1.8	90.8	6.2	95.1	2.2
avk	79.6	-13.4	95.1	4.6	93.0	-1.9	88.8	-8.7	93.6	-6.8
avn	95.1	1.1	97.9	0.3	95.7	2.9	95.1	3.3	96.4	0.5
avt	99.3	-0.7	99.8	0.1	99.6	-0.4	99.8	-0.4	99.8	-0.4
avu	97.7	1.7	98.3	-1.6	98.7	-0.6	97.3	-4.3	98.6	-2.0
awa	39.8	-1.2	71.9	32.3	77.2	-2.3	71.4	1.2	83.9	5.8
awb	98.6	-2.6	99.5	-1.1	99.4	0.0	99.1	1.2	99.3	-0.4
ayo	97.7	0.3	98.7	1.6	98.2	1.8	98.2	2.6	98.7	1.7
ayp	98.0	-0.7	98.4	-3.1	99.4	-0.4	98.2	-2.2	99.2	-0.5
ayr	97.1	3.8	97.9	0.1	96.6	1.7	95.1	4.1	98.5	1.2
azb	76.9	-17.5	81.9	-18.7	93.2	4.0	74.6	31.2	95.2	3.5
azd	93.3	-1.9	98.5	0.6	96.4	0.0	94.5	-3.0	96.5	-1.8
azg	99.1	0.6	99.8	-0.0	99.3	0.0	99.2	0.3	99.6	0.2
azj	93.8	0.7	93.7	-3.9	94.0	-1.0	91.5	-4.7	97.2	-0.5
azo	93.3	8.7	94.5	9.2	93.2	8.1	92.4	8.8	93.5	8.8
azz	88.1	-3.0	98.0	-3.1	97.8	-1.9	96.1	-3.0	97.4	-0.5
bak	85.7	1.4	91.0	-9.8	95.1	3.4	93.4	2.7	97.4	0.9
bam	69.1	19.7	77.7	32.9	67.6	-1.9	59.6	-3.4	75.1	3.4
ban	78.4	-11.3	92.9	1.8	91.6	2.3	83.1	-7.1	95.0	3.7
bao	97.4	0.7	99.2	1.2	97.2	-0.8	96.8	1.8	98.9	0.5
bap	86.4	-6.7	90.9	-1.3	94.5	4.4	91.2	-2.7	94.0	5.6
bar	36.2	-37.0	48.2	-16.7	60.1	14.7	39.4	44.5	57.7	52.1
bav	97.0	1.1	95.7	-3.7	94.9	-5.5	93.5	-8.8	94.0	-9.1
bba	98.8	-1.8	99.4	-1.0	99.5	-0.2	99.3	0.5	99.8	-0.2
bbb	96.2	0.8	99.7	-0.6	99.5	-0.5	99.7	0.2	99.7	0.0
bbc	63.0	18.6	75.0	22.0	72.9	13.1	69.7	1.5	77.6	13.4
bbj	96.7	4.2	98.3	2.5	96.0	2.8	95.8	1.1	97.7	1.3
bbk	97.1	3.1	98.6	2.0	96.0	0.9	96.4	3.4	96.7	3.7
bbo	97.3	-4.5	97.9	-4.0	99.4	-0.6	99.1	-0.4	99.0	-1.7
bbr	97.8	-2.5	99.4	-1.2	98.4	-1.1	98.6	0.8	99.2	-0.7
bch	96.3	-6.0	99.2	-1.3	98.7	-0.7	98.2	-0.5	98.9	-1.1
bci	94.0	4.7	95.2	8.0	94.2	-0.6	90.5	-5.9	95.3	-0.6
bcl	28.7	49.1	60.5	48.9	68.0	13.9	58.3	25.5	74.6	28.2
bcn	93.7	-0.1	96.9	4.1	93.2	-0.4	93.2	-2.1	95.4	-1.2
bco	96.6	6.4	96.0	3.9	96.1	5.7	96.5	5.8	97.0	4.8
bdd	96.4	2.1	98.9	-1.2	98.5	0.7	97.0	4.6	98.9	1.1
bdg	96.3	-5.2	99.0	-1.8	99.3	-0.7	99.5	-0.2	99.7	-0.3
bdq	94.9	-5.9	97.9	-1.8	98.3	0.5	97.5	-0.1	98.7	0.6
bdu	98.9	-1.6	99.4	-1.3	99.6	-0.3	99.5	0.0	99.7	-0.0
bdv	97.4	0.4	99.4	-0.2	99.3	1.0	99.3	0.6	99.5	0.4
bef	98.6	-0.9	99.3	-0.8	99.2	-0.4	98.5	-0.7	99.0	-0.9
beh	98.7	-2.1	99.5	-0.7	99.8	-0.2	99.9	0.1	100.0	-0.0
bel	94.4	-6.5	96.7	-4.3	97.2	-3.2	97.2	0.9	99.0	0.2
bem	67.6	-0.7	88.0	10.6	79.9	3.4	70.2	3.9	75.6	8.6
ben	74.7	-21.4	73.8	-31.0	94.5	-5.4	89.5	-12.3	97.3	-2.4
beo	96.2	-6.1	97.4	-4.8	97.7	1.7	97.0	4.8	98.6	1.8
bep	54.9	23.7	62.1	50.0	75.0	20.3	69.0	4.8	73.4	21.6
beq	84.5	-5.0	92.1	-3.1	93.2	-4.1	89.2	-4.7	92.3	-4.6
bet	91.1	-0.4	93.1	-1.2	96.7	-2.1	94.6	-2.2	97.2	-1.5
beu	95.9	-4.0	98.6	-2.3	98.5	-0.0	98.2	2.4	99.2	0.7
bez	78.6	24.1	88.1	14.5	93.9	-3.8	92.9	-3.0	96.5	-0.2
bfa	90.8	12.4	92.8	11.9	93.2	5.3	92.1	2.1	95.9	1.4
bfb	80.6	-7.1	98.3	-1.8	99.2	0.8	96.8	-0.0	99.0	-0.9
bfm	97.7	3.2	98.8	0.4	97.3	0.8	97.1	1.6	97.7	2.0
bfo	90.8	5.1	96.4	-3.7	96.0	-4.9	95.3	-2.0	98.0	-1.8
bfy	77.6	-1.6	93.7	6.8	92.0	-2.4	90.3	0.5	92.7	-0.9
bfz	81.6	4.0	97.8	-2.8	97.9	-1.5	95.2	-2.8	97.4	-3.0
bgc	72.1	12.3	96.3	-6.4	97.5	-2.2	94.9	-3.5	99.1	-0.1
bgg	85.9	13.7	86.9	9.7	88.6	6.0	79.5	-9.9	89.1	3.2
bgn	63.8	33.5	72.0	32.1	88.3	3.1	70.0	22.4	87.7	-1.3
bgq	68.0	-14.1	93.0	8.4	92.6	0.1	90.1	4.7	93.1	0.6
bgr	90.3	2.4	93.2	5.2	93.7	2.0	90.1	-0.1	93.9	1.9
bgs	96.3	-5.2	98.7	-1.2	99.3	0.7	98.7	1.6	99.2	1.1
bgt	95.5	-1.4	99.1	-0.9	98.2	-1.4	95.5	-3.5	98.3	-1.7
bgw	69.2	1.4	90.4	6.8	93.3	7.8	86.6	16.6	92.3	7.5
bgz	89.4	2.7	94.1	6.3	91.0	-2.0	91.5	3.7	93.5	2.2
bhd	85.1	11.5	95.8	3.9	95.1	0.3	94.1	5.7	96.1	-1.5
bhg	96.2	-2.7	99.3	0.0	98.1	-0.8	96.1	-0.8	98.6	-0.6
bhl	96.4	-3.9	97.0	-5.4	96.4	-2.3	96.4	1.3	98.2	1.4
bho	39.3	-17.1	72.5	-4.7	78.0	2.0	27.2	58.3	83.4	25.7
bhp	83.8	14.2	83.8	4.6	78.9	-10.7	83.0	2.2	85.1	1.7
bht	71.4	-7.0	97.0	-4.2	98.1	0.1	94.7	-2.1	98.3	-1.4
bhw	67.1	-16.1	58.3	-37.0	77.3	1.0	78.1	3.2	73.7	-9.5
bhz	71.7	-16.8	79.9	-31.7	87.0	-17.5	82.2	-8.1	85.6	-17.3
big	98.4	-1.9	99.2	-1.2	99.3	-0.1	99.3	0.3	99.4	0.1
bis	88.3	3.5	93.9	5.5	89.7	-1.3	86.2	-3.8	89.7	-7.2
biu	72.4	20.2	79.7	23.4	90.2	1.0	82.8	-10.1	91.2	2.1
bjk	98.8	0.4	99.8	-0.2	99.0	1.0	98.3	1.7	99.2	0.1
bjn	62.7	-31.0	84.2	-10.5	86.9	0.7	72.1	23.8	92.3	3.6
bjp	91.4	4.9	98.0	2.6	94.2	3.7	91.2	5.8	94.8	1.8
bjr	99.5	-1.0	98.7	-2.6	99.7	-0.3	99.9	0.2	100.0	0.0
bjv	98.5	-2.5	99.6	-0.4	98.7	-1.0	98.4	-1.4	99.4	-0.6
bjw	98.9	-1.7	93.8	-11.7	99.9	-0.2	99.5	0.5	99.8	0.3
bjz	98.5	-0.9	99.7	-0.2	99.1	0.3	97.9	2.8	99.5	0.5
bkd	98.5	-1.3	99.5	-0.7	99.3	-0.0	98.9	-0.1	99.7	0.1
bkh	94.4	1.8	98.1	1.5	95.7	2.4	92.7	0.9	94.5	-4.9
bki	98.5	0.2	99.6	-0.5	98.7	0.2	97.5	0.9	98.7	1.3
bkl	94.1	2.5	96.4	3.7	89.6	-7.9	91.8	1.3	92.7	-5.0
bkq	96.1	-6.2	98.5	-2.2	99.0	-0.1	99.0	-0.1	99.2	-0.5
bku	97.1	-4.8	98.9	-2.1	99.2	-0.6	99.4	0.3	99.6	-0.1
bkv	98.6	-2.8	99.6	-0.9	99.5	-0.5	99.5	-0.4	99.8	0.0
bky	97.8	2.0	96.9	3.9	96.4	3.9	97.1	4.2	98.7	2.1
blb	88.4	12.6	91.9	12.8	91.9	-2.4	90.0	-3.3	92.0	-3.9
blh	97.9	-4.1	98.9	-2.2	99.7	-0.3	99.7	-0.4	99.9	-0.1
blw	98.5	-1.7	99.8	-0.4	99.2	-0.7	99.6	0.3	99.8	0.0
blz	95.7	-1.3	98.0	-2.8	97.6	-2.3	96.0	-2.8	96.7	-3.9
bmh	98.1	-3.3	98.9	-1.9	98.2	-3.3	99.1	-1.3	99.6	-0.5
bmq	98.6	-2.5	99.1	-1.6	99.6	-0.6	99.7	-0.2	99.6	-0.7
bmr	99.4	-1.3	99.3	-1.4	99.9	-0.1	99.8	0.4	99.9	-0.1
bmu	99.3	-0.6	99.4	-0.7	99.7	0.3	99.4	-0.1	99.7	0.2
bmv	97.6	1.3	98.7	2.1	97.5	1.3	97.8	3.5	96.2	0.8
bni	92.5	0.6	97.8	-2.9	96.5	0.8	92.7	5.6	95.8	1.3
bnj	40.5	37.3	82.4	18.5	67.8	24.7	57.7	17.5	67.3	27.5
bnp	94.0	-8.6	94.0	-11.0	97.7	-2.7	95.7	-3.2	97.4	-3.4
bns	36.2	42.5	87.6	12.3	73.5	27.3	76.0	18.3	82.2	16.9
boa	99.2	-1.1	99.4	-1.3	99.5	-0.5	99.8	-0.2	99.8	-0.2
bod	63.1	-20.5	63.9	32.8	84.8	-18.4	69.3	-18.4	89.7	3.6
boj	99.3	-0.5	99.5	-1.0	99.5	-0.4	99.6	0.1	99.7	-0.4
bom	94.2	4.7	97.6	3.4	95.5	3.1	92.4	-2.7	96.2	1.9
bon	98.3	-0.7	98.8	-1.2	98.3	0.6	98.0	2.3	99.1	0.6
bos	45.1	-24.1	53.6	-34.4	50.1	-12.1	39.5	-2.1	54.4	-9.6
box	99.0	-0.1	99.5	-0.1	99.3	1.0	99.2	1.0	98.3	-1.7
bpe	95.8	3.7	97.9	2.9	97.0	0.0	94.9	-3.4	96.6	-2.9
bpr	71.8	-17.0	76.0	-29.5	76.0	-24.6	73.7	-16.2	76.7	-27.4
bps	62.2	12.8	61.9	39.3	65.2	29.5	65.1	18.7	63.8	36.5
bpy	89.3	1.7	92.7	-4.5	96.8	0.0	95.1	3.4	98.3	0.7
bqc	90.8	6.1	95.1	5.2	92.3	4.1	90.5	-3.4	90.7	0.3
bqj	95.2	-6.4	98.5	-1.6	98.2	-2.0	98.3	2.1	99.3	-0.0
bqp	79.8	-0.4	91.1	-3.5	82.0	12.5	79.7	3.6	82.6	14.8
bre	96.2	-4.6	96.4	-6.0	97.5	-2.4	97.1	-1.2	99.1	-1.3
bri	98.7	1.4	98.2	-0.2	98.1	2.2	97.9	0.9	97.9	-0.8
brv	98.1	3.7	98.0	3.2	98.4	3.2	98.4	3.0	97.2	0.2
bsc	88.6	18.8	90.7	15.6	94.4	5.9	93.8	2.7	93.8	7.1
bsj	96.5	-4.0	98.9	-1.7	98.2	-0.8	96.3	-0.4	97.6	-1.2
bsn	95.0	2.5	98.9	-1.7	98.8	-0.2	95.3	-0.2	98.4	1.8
bsp	92.2	-11.9	98.1	-3.4	96.2	-4.8	95.5	-6.4	96.4	-4.8
bsq	99.2	-0.7	97.4	-0.2	98.0	1.4	97.8	0.8	97.3	-0.5
bss	98.9	-1.7	99.1	-1.5	98.8	-0.1	98.0	1.0	99.4	0.4
bst	96.0	-0.4	78.6	-31.1	97.1	1.5	97.4	5.1	97.5	4.4
btd	89.5	6.6	94.5	9.4	93.0	2.8	87.9	-3.1	94.0	0.3
bth	91.8	12.9	92.8	7.3	91.9	7.3	91.4	6.4	94.4	5.7
btm	49.8	-31.6	77.3	-6.1	81.3	-2.8	71.6	13.9	88.5	-4.3
bts	71.7	20.4	78.8	25.5	77.7	25.4	75.6	20.2	79.4	29.0
btt	99.1	-1.4	99.8	-0.5	99.8	-0.3	99.6	0.1	99.8	-0.3
btx	65.4	21.8	77.1	33.2	73.7	21.6	67.2	19.7	77.3	6.7
bud	90.1	11.4	94.6	9.7	93.2	2.9	92.2	5.2	93.7	5.2
bug	94.2	7.0	96.9	4.8	94.1	0.9	92.1	-2.1	95.4	2.0
buk	98.5	-2.8	98.0	-3.6	99.0	-1.7	99.0	-0.8	99.3	-1.0
bul	85.4	-13.8	86.9	-19.3	97.0	0.9	93.6	-4.0	98.9	-0.9
bum	90.7	9.6	92.8	12.6	81.0	-17.1	89.0	-4.8	88.0	-12.6
bus	75.5	19.4	89.7	5.4	83.3	5.4	74.5	20.8	83.6	-4.1
buu	97.1	5.6	97.2	0.5	96.6	1.5	97.2	1.7	98.2	2.1
bvc	92.2	-5.0	99.6	-0.6	96.8	-0.4	91.8	5.8	98.6	0.3
bvd	90.6	0.3	99.1	-0.2	96.0	-2.3	89.1	-7.0	98.2	-0.6
bvr	99.6	-0.7	99.4	-1.1	99.5	0.3	99.6	0.5	99.7	-0.2
bwo	98.5	-0.1	99.1	-1.4	98.6	0.2	98.0	0.7	98.3	-0.7
bwq	97.1	-4.3	99.5	-1.0	99.4	-0.7	98.7	-0.6	99.7	-0.3
bwr	93.4	3.9	97.3	3.1	95.5	2.4	92.9	-2.3	96.7	-1.2
bxh	87.0	-9.7	97.5	-0.3	95.1	-1.8	91.3	-4.0	95.9	-2.6
bxk	74.1	4.1	91.1	7.7	89.0	3.1	84.6	0.7	90.1	4.7
bxr	92.6	-5.8	86.9	-22.3	97.7	-0.6	95.7	3.6	98.1	-0.7
byr	99.6	0.2	97.8	-3.5	99.2	0.5	99.2	0.7	99.5	0.7
byv	91.5	13.5	93.1	12.2	90.1	3.8	90.4	8.1	90.8	3.2
byx	97.7	-2.8	99.4	-0.4	99.0	1.0	99.1	1.7	99.1	1.6
bza	96.5	3.4	98.2	1.8	98.1	0.4	98.2	1.0	98.7	0.3
bzd	99.6	0.0	99.6	-0.5	99.7	-0.1	99.7	0.3	99.7	0.0
bzh	99.1	1.0	99.3	-0.4	98.4	0.2	98.1	2.1	98.7	1.9
bzi	88.0	-9.5	93.6	-5.7	97.7	-2.2	97.5	-0.8	98.4	-2.3
bzj	92.7	-5.6	98.8	0.2	97.2	1.2	92.2	6.0	97.8	0.7
bzw	92.6	10.8	94.1	7.3	92.0	6.7	91.0	4.6	93.0	4.6
caa	99.7	-0.4	99.9	0.0	99.6	-0.5	99.6	-0.0	99.8	-0.4
cab	96.9	-4.7	96.7	-5.8	97.4	-3.5	99.2	-0.6	99.0	-1.3
cac	99.4	-0.5	99.6	-0.4	99.6	0.0	99.3	-0.4	99.9	0.1
caf	75.3	-5.6	89.9	-14.6	88.9	-5.9	79.8	-15.7	88.6	-14.1
cag	97.9	0.8	98.7	1.6	95.8	-1.5	98.0	0.3	97.9	0.1
cak	93.2	-7.2	98.9	-0.5	98.2	0.3	97.1	-2.5	99.3	0.6
cao	98.2	-2.3	99.1	-1.1	97.0	-4.9	96.5	-4.3	98.1	-2.6
cap	99.1	-0.1	99.6	0.2	99.5	0.1	99.4	0.2	99.8	0.1
caq	97.5	1.5	98.0	2.5	97.1	-0.9	98.3	0.5	99.2	0.7
car	98.3	-3.0	99.5	-0.9	99.8	-0.1	99.4	-0.1	99.5	-0.7
cas	89.4	19.0	90.3	13.1	91.7	14.0	93.3	8.8	94.3	9.4
cat	65.4	-17.6	85.8	-5.7	84.7	-5.1	67.9	-20.4	89.4	-6.1
cav	97.1	-3.1	96.2	-6.0	98.5	0.6	98.1	1.8	97.5	-1.5
cax	98.6	0.9	99.3	-0.3	98.8	-0.3	98.8	1.3	99.0	0.7
cbc	96.0	4.2	99.4	-1.0	98.5	-0.2	97.1	0.8	99.0	1.0
cbi	98.5	-0.7	99.6	-0.7	98.5	1.4	97.6	4.0	98.6	1.7
cbk	75.5	9.3	84.5	8.4	79.0	-6.9	78.1	3.8	79.6	-14.9
cbl	93.3	1.3	97.8	-3.5	98.4	-1.5	96.5	2.4	99.1	0.2
cbr	93.9	9.9	97.4	-0.4	91.2	15.9	89.7	18.4	97.2	5.2
cbs	97.1	-1.5	98.8	-1.0	97.2	-0.8	96.8	2.3	97.4	0.8
cbt	99.1	-1.0	94.0	-10.4	99.5	0.2	99.4	0.1	99.5	0.0
cbu	92.8	-5.8	98.1	-2.7	93.0	-4.1	88.3	6.3	93.6	0.6
cbv	99.3	0.1	99.4	0.3	99.4	0.4	99.4	0.9	99.5	0.4
cce	85.1	8.5	95.1	7.9	92.9	2.6	85.4	-2.1	94.0	2.0
cco	99.9	0.1	99.8	-0.1	99.7	0.4	99.8	0.5	99.9	0.2
ccp	52.9	60.8	66.0	49.9	97.9	0.5	96.6	-4.0	98.2	0.2
cdf	96.5	-5.4	98.4	-2.9	99.0	-1.1	98.0	-2.4	99.0	-0.3
cdh	56.3	-2.9	95.8	-6.0	97.6	0.1	92.3	-1.7	98.0	-1.2
cdj	77.5	16.1	97.5	1.0	94.3	8.3	92.2	8.1	98.0	3.3
cdo	70.4	34.2	68.3	28.6	88.4	-3.7	91.8	-9.1	93.0	-7.9
ceb	78.5	-7.7	92.9	-3.4	92.8	-3.7	83.4	12.5	96.2	5.0
ceg	99.8	-0.4	99.7	-0.5	99.8	-0.4	99.9	0.2	100.0	0.0
cek	93.7	-1.0	97.3	-1.4	89.7	-12.0	94.9	0.6	96.3	-1.5
cen	86.4	19.3	94.9	8.0	92.4	6.8	89.4	1.2	93.9	2.5
ces	94.1	2.1	95.4	-1.5	92.9	-2.1	91.0	-2.3	95.9	-2.5
cfm	71.5	21.0	79.6	22.7	77.1	-1.4	68.4	9.1	77.0	-16.3
cgc	93.3	4.8	98.6	-2.0	97.5	-1.0	95.0	-2.5	97.8	-3.0
cha	97.1	-2.6	99.2	-0.8	98.6	-0.3	98.8	1.9	99.3	0.2
chd	98.7	-0.7	98.5	-2.0	98.0	-0.3	98.4	0.0	98.8	-0.4
che	88.2	3.9	75.5	-33.5	93.8	3.2	85.4	0.3	94.9	2.5
chf	98.6	0.6	99.4	-0.6	99.1	0.1	98.3	-1.2	99.4	-0.6
chj	99.6	-0.4	97.4	-4.7	97.6	-3.5	99.4	0.0	97.5	-4.4
chk	95.7	4.9	96.9	3.5	96.0	5.9	96.0	6.5	97.9	3.3
chq	99.2	0.0	99.5	-0.5	99.2	0.1	99.7	0.5	99.8	0.2
chr	100.0	0.0	93.2	-12.8	99.3	-1.1	100.0	0.0	100.0	0.0
chv	85.6	10.2	86.1	0.4	91.7	1.4	92.4	9.7	95.6	2.5
chy	99.4	0.7	95.9	-6.8	99.2	0.0	99.5	0.6	99.5	0.4
chz	99.9	0.2	99.9	-0.1	99.8	0.1	99.8	-0.3	100.0	-0.1
cja	99.2	-1.6	98.6	-2.7	100.0	0.0	99.9	0.1	100.0	-0.1
cjk	76.9	14.3	90.6	10.6	89.8	4.2	83.5	7.0	91.7	4.8
cjo	68.8	2.7	84.2	-11.4	86.8	-2.4	79.8	20.4	90.8	7.1
cjp	99.4	-0.6	99.7	-0.5	99.8	0.3	99.9	0.0	99.9	-0.3
cjv	98.5	-0.9	99.7	-0.7	99.8	-0.3	99.4	-0.5	99.2	-1.3
ckb	95.0	-7.2	94.8	-8.4	98.2	-2.5	97.7	-3.0	98.7	-1.7
cla	90.9	7.8	96.3	3.5	93.2	0.4	89.1	-4.7	94.3	0.0
cle	99.5	0.8	99.8	-0.1	99.5	1.0	99.4	1.0	99.6	0.6
clt	92.5	10.4	95.6	5.8	93.6	1.1	90.9	-3.2	94.2	-1.8
clu	87.8	-5.9	95.8	-5.6	96.0	-1.4	93.1	-4.7	96.8	-1.8
cly	95.5	-0.3	99.4	0.0	98.0	-2.1	98.4	2.3	99.3	0.9
cme	98.5	0.1	98.9	-1.7	98.1	1.7	97.8	1.3	98.6	0.7
cmi	94.2	4.7	92.4	-2.3	90.7	-10.9	89.8	-11.5	92.5	-11.0
cmn	51.9	-18.6	44.3	8.5	57.0	-1.4	74.4	8.5	83.4	9.7
cmr	96.8	2.7	98.2	2.3	94.7	4.5	93.4	-1.3	96.8	2.1
cnh	66.4	-10.5	77.2	31.0	74.7	8.3	66.6	-4.9	79.4	20.5
cni	96.2	-4.5	99.1	-0.9	98.4	-0.6	96.6	0.6	98.0	-1.9
cnk	93.7	9.3	94.9	8.9	93.9	5.2	89.5	-4.1	94.7	0.5
cnl	99.6	0.0	99.7	0.1	99.7	0.5	99.6	0.5	99.5	0.8
cnr	56.4	6.9	68.8	-5.5	64.3	10.5	11.3	32.7	57.0	32.3
cns	91.0	-4.2	96.1	1.3	95.3	3.4	95.1	5.7	96.1	3.2
cnt	99.4	0.7	99.2	-0.7	98.9	1.5	99.6	0.3	99.0	0.8
cnw	65.6	12.8	79.3	26.8	78.1	7.3	67.6	11.4	82.2	-9.2
coe	99.6	0.4	98.9	-1.9	99.7	-0.2	99.7	0.5	99.7	0.3
cof	98.1	-2.9	99.2	-1.2	99.3	-0.2	98.6	-0.4	99.6	-0.2
cok	95.0	8.9	96.5	6.6	95.2	7.4	95.8	7.0	96.4	5.3
con	97.0	5.3	96.9	3.3	98.1	2.0	99.1	1.5	99.1	0.4
cop	100.0	0.0	62.2	-35.3	99.7	-0.4	100.0	0.0	100.0	0.0
cor	79.1	-5.2	87.9	10.5	89.8	-6.5	89.1	-2.3	93.7	-1.1
cos	71.3	25.2	72.9	38.2	76.1	4.5	73.8	-5.5	86.0	7.2
cot	97.2	0.2	99.0	-1.0	98.7	0.0	96.4	-3.4	98.7	-0.4
cox	95.9	5.9	95.8	1.4	95.9	3.0	96.6	5.6	97.0	2.7
cpa	99.3	-1.2	99.6	-0.7	99.9	-0.1	99.5	0.1	99.7	0.1
cpb	80.3	-4.8	94.1	-4.2	94.4	-2.7	84.1	10.0	94.7	-1.6
cpc	83.6	-11.9	94.8	-7.3	96.0	-3.6	93.1	-7.0	96.5	-3.3
cpu	89.1	-8.3	97.0	-3.7	96.8	-1.3	90.3	-5.9	96.5	-0.7
cpy	60.4	10.1	84.8	10.3	85.5	5.8	77.9	-16.0	88.5	0.9
crh	3.1	41.7	51.3	43.3	86.3	2.6	72.6	13.3	91.7	6.9
crk	76.8	37.1	66.2	48.9	96.4	5.4	98.9	1.3	98.9	1.0
crm	94.2	-9.7	67.6	-48.7	99.3	-0.3	99.5	0.4	99.6	0.2
crn	98.8	-0.7	99.8	-0.1	99.1	-0.6	99.2	-0.9	99.6	-0.2
cro	98.6	-2.2	97.7	-4.0	98.2	-2.4	98.7	-0.3	97.5	-3.9
crq	99.5	-1.0	99.9	-0.3	99.8	-0.5	99.5	-0.7	100.0	-0.1
crs	85.8	11.1	94.9	-0.4	92.9	2.3	86.1	3.2	89.7	6.9
crt	91.7	12.4	91.9	11.7	84.1	-8.9	85.7	-7.2	91.3	-3.8
crx	73.1	6.3	88.9	14.7	87.4	7.0	73.8	21.4	85.6	18.6
csb	84.2	-15.9	86.8	-14.6	90.2	-5.5	92.9	3.3	95.1	2.2
csk	85.6	7.6	93.2	7.7	89.9	-6.2	86.5	-11.9	92.0	-6.7
cso	99.9	-0.1	99.9	0.0	99.6	0.1	99.2	-0.5	99.6	-0.3
csw	94.9	1.6	95.6	1.0	97.1	-0.6	97.7	-1.0	98.9	-0.1
csy	84.7	-18.5	94.9	-9.6	96.4	-4.2	93.9	-6.6	98.2	-0.4
cta	99.4	-1.2	99.4	-1.3	99.9	0.2	99.8	-0.1	99.9	0.0
ctd	70.1	16.2	83.6	22.8	82.5	0.9	75.3	-1.9	83.3	19.6
ctg	95.1	-6.3	96.6	-5.5	99.6	0.5	98.9	-0.4	99.6	-0.1
cth	91.0	-14.3	95.6	-6.5	95.9	0.0	95.2	-0.2	98.0	1.1

Table 5: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
cto	99.4	-0.7	99.1	-1.7	99.6	-0.5	99.5	-0.1	99.8	0.2
ctp	99.1	-1.5	100.0	-0.1	99.7	-0.1	99.6	0.0	99.9	-0.0
ctu	98.5	0.3	98.6	-1.1	98.7	1.2	98.5	2.4	99.4	1.0
cub	99.5	-0.3	98.2	-3.6	99.7	0.2	99.4	0.1	99.8	0.1
cuc	99.7	-0.3	99.6	-0.6	99.6	-0.3	99.4	-0.9	99.7	-0.5
cui	97.5	0.1	99.0	-0.2	99.0	0.6	97.5	3.6	99.0	-0.1
cuk	95.6	-4.6	96.0	-7.4	91.8	-12.9	90.7	-14.7	91.7	-14.3
cul	99.1	-1.0	96.8	-6.2	99.7	-0.5	99.9	-0.1	99.8	0.0
cut	99.6	-0.2	99.8	-0.2	99.6	0.6	99.3	0.0	99.6	0.7
cuv	91.5	8.0	94.9	7.4	94.7	-3.4	89.4	-8.2	93.6	-4.9
cux	99.8	-0.4	99.5	-1.0	99.9	-0.2	100.0	0.0	100.0	0.0
cwa	89.4	5.0	93.2	4.0	90.5	-6.0	88.2	-8.7	87.2	-15.4
cwe	34.8	-10.5	70.7	17.1	73.0	10.7	50.5	16.2	61.5	30.6
cwt	97.3	-4.7	99.1	-1.8	99.4	-1.0	99.5	-0.2	99.7	0.0
cya	95.6	0.3	99.4	-0.3	97.9	0.8	98.6	-2.2	99.4	-1.1
cym	93.2	-2.2	88.6	-16.4	94.4	-6.8	93.9	-4.3	97.6	-0.2
czt	92.1	9.8	75.3	-25.0	91.8	3.9	87.8	1.3	92.6	-1.4
daa	95.4	-5.6	98.7	-1.9	98.3	-0.8	96.5	-2.0	98.3	-1.5
dad	96.9	3.3	98.7	0.0	96.6	1.1	96.6	1.1	98.1	0.7
dag	90.2	-3.0	94.6	-1.3	93.0	2.9	90.7	0.5	94.0	4.8
dah	98.0	-2.8	99.1	-1.3	99.2	-0.1	99.1	0.5	99.4	-0.2
dak	98.0	-0.4	98.5	-1.2	98.4	1.1	98.2	2.0	97.7	-0.3
dan	78.8	-9.3	91.3	-12.3	93.7	0.3	83.7	0.3	98.5	1.4
dao	59.4	-15.1	69.5	-46.4	66.0	-15.6	58.4	-3.2	70.5	-35.8
dav	80.6	11.4	89.6	8.4	87.2	8.3	82.7	11.3	88.6	11.6
ddn	96.6	-5.5	99.2	-1.6	99.4	-0.3	99.2	-0.2	99.7	-0.2
ded	98.7	-1.4	99.4	-0.7	99.3	0.5	98.5	2.0	99.4	0.8
des	98.7	-2.1	99.3	-1.5	99.6	-0.9	99.2	-0.9	99.6	-0.3
deu	79.2	-25.8	85.7	-20.9	90.7	-10.3	78.7	-26.0	93.5	-8.7
dga	90.7	6.7	97.0	4.5	93.2	3.8	91.6	0.9	91.8	-4.5
dgc	90.8	-6.3	99.3	-0.5	98.1	-0.8	97.9	0.2	99.2	-0.6
dgi	91.0	1.7	96.4	4.0	94.3	2.0	93.3	0.2	96.1	0.7
dgk	99.1	-1.7	99.0	-1.9	99.8	-0.2	99.8	0.1	99.9	0.1
dgo	70.4	-12.7	96.4	-1.9	97.9	1.6	95.3	1.2	98.3	0.7
dgr	99.9	0.0	99.5	-0.5	99.1	0.4	99.6	0.4	99.7	0.2
dgz	96.7	-2.9	97.2	-4.7	97.6	-1.8	95.4	-3.7	98.1	-2.0
dhd	71.4	18.3	86.9	18.2	88.8	7.9	85.3	3.9	89.3	8.0
dhi	95.8	-1.6	97.7	-0.8	98.6	1.8	98.4	2.6	98.8	1.9
dhm	56.4	0.5	79.4	18.0	78.9	-3.4	72.6	-10.7	79.8	-8.6
dhn	52.4	19.1	76.5	9.4	79.6	-6.0	72.1	8.6	76.9	13.8
did	98.8	2.2	99.1	1.4	97.7	1.1	98.4	0.0	98.2	-1.2
dif	98.1	-3.1	98.7	-1.6	99.2	-0.4	99.4	0.1	99.2	-0.4
dij	30.2	16.5	1.0	11.1	40.9	12.9	41.9	-2.8	28.7	22.9
dik	80.4	-3.6	88.6	-2.0	89.9	-2.4	87.4	3.7	92.9	1.9
dip	76.5	16.5	89.9	5.8	90.5	5.0	87.5	1.9	92.0	1.6
diq	46.7	5.4	44.2	-5.9	67.8	13.4	17.8	37.9	3.5	46.4
dis	85.9	15.5	88.4	19.1	86.5	12.1	83.6	13.2	87.0	13.0
div	100.0	-0.1	99.9	0.2	100.0	0.1	100.0	0.0	100.0	-0.1
dje	93.3	-4.0	97.9	-1.2	97.1	1.1	99.1	0.6	99.5	-0.1
djk	89.1	1.2	97.1	2.4	94.9	3.6	92.5	1.3	95.6	4.1
djr	94.0	-4.8	98.8	-0.6	97.7	0.4	96.1	4.1	98.7	0.7
dks	81.1	19.6	76.1	36.3	89.4	-2.0	83.3	-15.4	89.8	-7.1
dln	94.9	-3.3	95.8	-7.4	98.5	-1.0	95.8	-3.2	99.3	0.0
dna	82.6	8.6	89.5	15.7	83.8	2.8	80.5	1.8	80.1	-4.5
dni	87.1	18.2	88.2	12.7	85.7	11.6	77.8	-10.2	85.9	9.6
dnt	86.2	14.9	89.5	15.7	83.1	2.6	79.4	-1.4	86.6	12.1
dnw	44.9	19.0	60.6	26.5	55.6	1.1	49.4	7.9	59.5	-5.1
dob	97.5	-3.1	99.2	-1.4	98.8	-0.9	96.0	-6.5	99.0	-0.2
dop	99.2	0.6	97.7	-4.1	97.7	3.9	97.8	4.2	98.5	2.9
dos	99.5	-0.9	99.0	-2.0	99.6	-0.4	99.4	-0.4	99.6	-0.5
dov	68.0	-3.9	88.2	-14.6	87.2	-10.3	78.2	-9.1	83.5	-14.0
dow	92.7	5.8	94.6	4.7	95.9	-0.3	94.6	-1.8	95.7	-0.9
dru	81.4	16.9	93.6	8.6	93.5	2.7	91.7	-0.8	93.3	-1.1
dsb	56.5	-7.5	55.8	9.9	66.8	5.3	56.6	-2.8	80.2	-0.4
dsh	97.5	3.4	97.8	1.6	96.7	0.7	96.7	-0.6	97.6	-1.3
dso	91.5	-13.5	99.5	-0.7	99.4	-0.1	99.1	-1.2	99.6	-0.3
dtb	80.6	13.3	89.6	13.9	85.7	-9.5	81.1	-21.1	83.6	-20.8
dtp	78.0	22.4	85.5	23.0	72.0	-16.0	82.1	22.0	73.4	-15.8
dty	42.4	-12.1	71.6	-0.4	79.7	-0.6	63.5	41.9	81.4	-0.1
dub	30.2	25.9	39.9	49.2	39.7	68.0	34.3	35.8	37.9	50.8
due	88.4	-17.6	97.0	-4.9	97.8	-3.1	97.5	-2.5	99.4	-1.0
dug	75.1	14.3	83.4	9.4	83.0	1.7	70.9	-10.3	80.3	-5.1
duo	94.5	-2.9	99.2	-1.1	98.3	-2.6	98.2	-0.1	99.1	-0.7
dur	90.8	15.9	91.2	15.6	95.1	-2.0	94.9	-1.8	95.6	-3.7
dwr	45.8	55.0	51.9	63.1	87.0	8.8	78.5	11.6	87.6	5.7
dww	96.2	-4.0	98.4	-1.5	98.1	-0.5	97.0	4.0	98.4	1.3
dyi	92.2	-6.3	97.7	2.3	98.5	0.3	98.3	0.2	99.2	0.3
dyo	97.0	-4.7	98.5	-2.9	98.6	-2.0	99.0	-1.1	99.8	-0.2
dzo	78.9	9.0	89.1	16.0	97.0	-0.5	93.1	-7.0	97.5	-3.3
ebk	94.7	1.7	99.4	-0.7	98.7	-0.6	98.2	1.2	99.0	-0.1
ebr	94.3	4.8	96.4	5.6	95.6	-4.7	94.5	-5.0	95.9	-4.5
ebu	83.0	6.8	91.7	5.5	91.5	2.9	86.3	9.9	92.0	2.8
efi	94.3	3.9	96.6	-1.1	96.7	2.8	96.9	4.7	98.0	2.8
egl	80.7	0.5	86.6	3.3	79.8	-5.8	80.5	19.0	90.3	10.8
ego	98.0	2.0	99.0	-0.5	97.8	0.9	98.3	1.1	98.7	0.8
eip	96.0	0.9	96.8	4.2	98.6	0.1	97.3	4.1	99.5	0.6
eka	96.4	4.2	98.6	1.1	96.2	-1.3	95.8	-0.7	97.1	-0.5
ekg	93.4	4.4	97.1	2.4	94.3	-2.1	92.2	-4.1	95.9	-0.6
ekk	91.9	-12.8	94.3	-10.3	96.4	-5.5	90.2	-15.8	91.8	-14.4
eko	90.0	-7.0	97.5	-2.6	95.4	-3.0	91.9	-4.4	95.8	-2.5
eky	96.2	6.4	98.4	1.1	99.3	0.7	99.5	0.0	99.9	-0.0
ell	83.3	-0.4	77.9	-29.7	92.5	-2.3	88.6	-6.6	96.7	-3.6
emi	96.7	1.3	98.9	-0.1	97.7	0.9	96.3	1.9	98.3	0.4
emp	97.1	3.8	98.2	-1.7	97.4	3.2	96.8	5.8	99.8	0.1
ena	76.4	32.6	77.0	31.0	84.8	13.3	87.5	9.7	90.0	3.6
enb	98.7	-1.8	99.6	-0.9	99.6	-0.4	99.5	-0.2	99.9	0.1
eng	60.2	-33.0	69.5	-24.9	72.8	-12.2	66.7	-3.0	85.3	-3.3
enl	92.8	6.6	92.7	6.3	92.2	4.8	92.2	0.6	94.7	-0.7
enm	91.4	-1.1	97.8	-3.2	96.7	2.1	94.0	-0.5	99.3	-0.0
enx	96.9	5.3	97.5	4.4	95.7	3.6	94.1	-0.5	96.0	2.2
epo	82.0	-12.5	83.1	-18.2	88.0	-3.4	80.5	-19.9	93.2	-0.6
erg	98.4	0.5	98.4	-0.9	97.9	0.5	99.2	0.8	99.1	1.0
eri	96.8	-1.6	99.4	-0.4	98.1	1.1	97.0	3.7	98.8	0.7
ese	98.7	-1.9	99.6	-0.9	99.9	-0.1	99.5	0.7	99.6	0.5
esg	82.5	-2.6	96.4	1.2	93.6	-7.3	91.0	-9.2	94.4	-6.2
esi	78.3	16.9	85.8	7.7	88.1	-2.3	78.3	10.2	89.8	-0.7
esk	83.3	-10.5	91.0	-10.5	90.6	-1.4	81.8	-8.4	91.7	1.9
ess	64.4	50.5	94.8	9.2	98.6	0.7	98.8	0.9	99.3	0.6
esu	81.9	-25.8	97.2	-1.1	95.0	-2.5	93.2	-3.6	96.4	-2.4
etn	61.3	5.4	82.4	18.9	84.9	8.4	77.0	8.7	78.1	25.1
etr	98.3	-2.5	97.6	-4.4	98.2	-1.5	99.3	0.9	99.6	0.6
etu	89.0	-0.5	82.2	-20.8	90.4	-6.9	90.7	-7.6	94.5	-3.2
etx	86.5	19.0	86.4	22.6	91.7	9.2	91.9	5.9	95.8	3.4
eus	84.1	-13.7	92.8	-5.9	90.3	-9.2	87.9	-10.0	95.8	-2.2
ewe	98.6	-1.1	99.5	-0.6	98.5	-0.8	98.1	-1.3	99.3	0.3
ewo	95.1	6.6	96.5	4.4	94.4	-2.3	94.2	-2.4	96.2	-1.5
ext	79.4	-2.7	90.6	-10.1	90.3	-4.5	83.8	15.2	95.5	3.7
eza	83.2	-6.1	95.2	2.8	93.4	-0.8	88.9	-9.7	94.4	-2.5
faa	97.7	-4.3	97.6	-4.6	99.5	-0.8	99.8	0.3	99.8	-0.2
fai	97.1	-4.0	98.7	-1.6	98.4	-0.3	98.4	1.4	98.8	0.3
fal	99.0	-1.4	99.3	-0.9	99.7	-0.1	99.6	-0.1	99.6	-0.6
fan	96.5	-0.8	98.5	0.1	97.3	1.9	96.5	3.2	97.4	0.5
fao	84.7	13.9	92.6	-0.5	90.4	0.8	88.6	5.0	95.3	3.5
far	94.2	-3.7	97.5	-4.2	97.5	-1.7	93.7	-3.4	98.2	-0.8
fat	89.4	11.2	95.2	4.9	92.8	5.8	88.7	-1.1	92.9	-0.2
ffm	67.3	-4.3	87.1	-1.1	85.9	-4.4	70.0	4.0	89.2	2.0
fij	88.9	3.3	93.8	6.9	91.7	-0.2	88.9	-6.1	93.5	0.7
fil	46.8	-11.6	69.2	0.4	65.0	-16.9	56.8	-32.1	69.8	-28.3
fin	71.9	6.1	84.1	5.7	85.4	-10.5	74.3	-21.3	88.9	-13.1
fip	69.8	21.8	87.2	15.6	72.1	17.6	62.5	19.6	77.4	6.0
fit	74.7	-11.2	87.9	-11.4	88.3	-6.3	75.1	6.4	90.9	-3.4
fmu	69.2	-1.6	84.5	11.0	85.2	1.3	80.9	-4.8	88.4	3.8
fon	94.5	5.7	81.4	-24.3	95.5	5.0	95.1	1.7	97.0	2.4
for	96.1	-5.6	99.0	-0.3	98.8	-0.3	99.1	0.6	99.3	-0.1
fra	82.7	-3.9	89.2	-8.0	89.1	-5.4	69.9	-31.7	90.0	-13.9
frp	71.7	-1.7	84.6	2.5	82.5	-3.8	71.8	-2.6	79.6	-16.3
frr	70.4	-5.4	90.4	0.5	89.3	1.4	66.3	22.8	94.5	1.7
fry	88.0	-16.4	83.7	-26.5	89.9	-14.9	81.6	-20.1	88.6	-18.9
fub	61.8	46.7	15.7	80.1	83.9	7.6	75.8	15.5	86.0	10.3
fud	89.0	-6.3	94.9	3.4	93.5	3.8	92.5	-0.9	94.5	2.0
fue	69.0	-7.4	89.9	-14.3	90.9	-5.5	78.9	-15.5	93.0	-3.8
fuf	78.9	1.9	94.0	-4.9	92.9	-0.7	82.9	-4.6	94.7	2.3
fuh	61.0	-2.3	83.8	-1.5	84.8	0.2	68.7	17.2	89.0	0.3
fuq	47.2	28.9	73.1	27.7	79.7	4.8	63.4	-16.6	82.7	-3.5
fur	77.0	14.3	87.5	12.5	81.4	8.4	67.9	35.1	85.3	20.9
fut	95.9	2.2	96.4	1.1	99.3	-1.1	99.1	-0.4	99.6	-0.2
fuv	55.0	-15.0	71.8	-15.3	83.2	-9.8	73.6	-0.3	87.9	-7.3
gaa	84.7	11.7	84.9	7.3	83.8	-1.4	81.2	-6.5	87.3	0.2
gab	97.0	5.7	97.3	3.9	94.3	-3.7	95.7	-2.6	98.2	-1.2
gag	3.5	41.4	69.4	38.7	91.6	-2.4	85.9	6.9	91.8	-6.2
gah	98.5	-2.8	99.2	-1.7	99.7	-0.2	99.7	0.2	99.6	-0.2
gai	98.6	0.7	99.5	-0.1	98.7	0.3	98.3	2.6	98.9	1.4
gam	98.3	-0.4	98.9	0.0	98.8	1.2	98.2	1.2	98.9	1.1
gan	60.9	-35.1	55.6	-38.4	69.8	-26.0	76.6	-17.9	77.4	-28.4
gaq	92.8	-0.2	98.1	-1.0	99.0	1.2	98.0	1.5	99.0	-0.6
gas	91.6	7.8	97.2	0.9	97.8	3.8	96.1	2.0	98.7	-0.1
gau	88.3	0.3	71.4	-36.1	97.5	3.3	97.2	0.5	98.5	0.5
gaw	99.3	-1.5	99.3	-1.0	99.5	0.0	99.8	0.5	99.8	-0.2
gax	89.8	9.0	93.3	9.2	93.2	3.6	91.8	0.4	95.8	-0.3
gaz	79.1	-16.2	91.0	3.1	92.9	-1.2	83.1	16.1	94.7	1.2
gbi	87.4	4.8	96.5	4.5	93.8	-2.5	91.4	0.7	95.9	-2.2
gbk	72.0	-6.4	97.0	-2.8	96.6	-1.0	92.9	9.1	98.0	0.3
gbl	68.7	-10.0	85.2	-14.4	90.0	-8.6	82.5	-5.4	85.5	-17.1
gbo	97.0	0.8	97.7	-0.3	94.4	-4.9	93.0	-8.2	96.3	-2.9
gbp	99.4	-1.2	99.8	-0.5	99.4	-0.4	99.9	-0.0	99.8	-0.1
gbr	96.5	1.7	96.8	-2.8	96.7	0.3	95.3	-0.0	96.5	0.4
gcr	93.5	3.0	95.7	-0.6	95.3	0.2	89.8	-6.2	95.6	-2.2
gdg	97.8	-1.4	97.6	-4.6	99.0	-1.6	99.3	0.3	99.6	-0.1
gdn	98.1	0.6	99.6	-0.9	99.0	-0.3	98.4	1.9	99.4	-0.3
gdr	98.9	1.6	99.3	0.6	98.5	1.3	98.8	1.3	99.0	1.0
geb	99.3	-0.7	99.3	-1.1	99.6	-0.1	99.4	0.6	99.7	0.1
gef	98.7	-0.7	99.2	-1.3	99.5	-0.5	99.0	0.6	99.7	0.3
gej	98.0	0.3	98.7	0.7	97.4	3.7	97.4	0.3	97.8	0.8
gez	87.6	2.3	28.8	56.3	98.5	-1.0	97.5	3.2	99.0	-1.0
gfk	95.0	-3.1	99.5	0.4	97.8	-1.8	94.5	-4.3	98.1	-2.9
ggw	82.4	11.7	85.9	14.8	94.7	-3.2	94.3	-1.3	95.8	-2.4
ghe	79.3	10.1	88.6	12.4	91.8	4.0	90.3	5.7	92.1	5.2
ghs	98.1	-0.0	99.3	-0.7	98.7	1.4	98.4	2.5	98.7	1.4
gid	90.5	-6.4	96.4	0.3	92.7	-4.6	90.5	-7.5	91.4	-10.9
gil	90.1	10.1	94.4	7.7	89.9	4.0	89.4	2.2	92.1	4.9
giz	92.4	7.3	96.4	6.1	93.7	7.9	94.2	4.6	97.8	4.0
gjn	95.9	-2.8	98.5	2.3	97.7	0.6	97.2	2.0	97.4	0.5
gkn	99.0	-2.1	99.4	-1.2	99.8	-0.1	99.8	-0.2	99.9	0.1
gkp	96.8	5.1	98.2	2.4	95.9	5.7	96.6	3.2	97.5	3.9
gla	97.2	-1.2	97.5	-3.0	97.1	-0.8	96.2	-0.2	98.7	0.4
gle	92.8	0.2	94.1	0.7	93.2	-2.5	89.5	-6.9	95.7	-2.1
glg	78.8	1.4	89.8	2.2	89.6	2.4	77.5	3.1	93.3	6.9
glk	50.9	0.5	62.6	30.3	74.1	-20.6	67.7	-16.5	75.2	-24.2
glv	79.4	19.4	81.1	10.4	79.9	6.3	80.6	13.9	88.6	14.2
gma	85.1	-14.7	91.2	-5.5	90.5	-9.0	92.1	-6.2	92.6	-9.8
gmv	56.4	49.6	64.8	48.0	94.0	-2.2	88.8	-5.9	94.5	-2.5
gna	92.3	13.6	96.7	6.0	94.9	-2.1	93.7	-4.2	94.9	-4.9
gnb	63.6	20.1	79.6	29.8	77.9	2.5	72.4	6.1	77.6	-0.1
gnd	95.5	5.4	97.6	3.6	95.2	-0.2	94.2	-0.6	95.7	-1.9
gng	97.0	1.5	99.2	-0.6	98.2	1.8	96.9	4.7	97.8	2.8
gnn	94.4	4.4	99.1	-0.4	98.4	0.3	96.2	-2.4	98.9	-0.0
gnw	90.1	8.2	96.5	2.2	95.1	2.5	90.0	0.6	95.2	-3.8
goa	92.4	7.9	94.7	8.9	94.5	1.0	94.0	1.1	95.4	-2.7
god	86.0	5.4	92.6	13.4	96.2	-1.3	94.5	-2.6	97.1	-0.8
gof	48.3	43.6	52.4	62.0	88.5	-5.7	80.4	-11.2	88.6	-6.1
gog	82.7	-3.4	94.2	-2.1	93.0	0.2	90.4	-3.5	94.4	2.6
gok	74.2	15.7	98.6	-2.0	99.0	-0.4	96.2	3.9	99.8	0.2
gol	97.9	-0.7	98.9	-0.0	98.6	-0.4	99.1	0.6	99.5	-0.2
gom	0.0	0.0	60.4	53.0	89.7	-4.0	84.7	1.9	95.6	3.8
gor	72.1	-1.0	86.5	8.2	90.8	-5.5	89.9	0.6	95.0	0.1
gos	72.9	0.3	58.4	52.6	86.9	-6.2	75.6	-22.7	90.3	-13.5
got	66.8	49.8	64.6	52.2	96.1	-3.7	98.4	-0.8	99.2	0.0
gqr	97.3	4.5	98.1	2.2	95.7	-0.7	97.2	1.4	98.3	1.7
grc	84.4	-1.1	65.3	39.5	93.6	-0.4	87.8	6.2	96.8	3.2
grt	92.3	-8.3	96.0	-3.5	98.3	2.7	97.0	-0.2	99.1	1.3
gso	93.6	1.2	97.6	3.8	89.2	-13.0	93.2	-1.8	94.7	-5.5
gsw	76.4	-21.7	76.4	-34.7	88.5	-12.4	81.4	8.7	94.6	-0.8
gub	97.8	0.9	98.6	0.5	98.1	1.4	98.4	2.5	98.8	1.4
guc	99.1	-1.3	99.3	-1.1	99.8	-0.1	99.5	0.0	99.9	0.1
gud	94.9	-0.1	96.4	0.9	94.7	-0.9	92.6	-4.0	93.6	-4.8
gug	77.4	33.6	80.3	32.2	90.3	4.7	86.8	-10.6	92.6	-2.7
guh	96.6	-5.8	98.3	-3.4	99.1	-1.7	99.4	0.0	99.9	-0.3
gui	90.0	-7.9	96.5	-1.9	95.0	-3.5	89.3	-4.2	95.6	-0.3
guj	64.4	-38.5	69.1	-32.8	73.7	-39.1	69.7	-39.9	73.9	-40.0
guk	98.8	-1.0	99.2	-0.8	99.8	-0.0	99.7	0.4	99.9	0.0
gul	97.4	-0.6	99.1	-1.0	97.4	-1.8	95.7	-2.7	99.1	-0.8
gum	98.0	-0.5	99.2	-1.5	98.7	0.9	98.9	1.9	99.3	0.7
gun	93.6	-3.4	97.0	-4.4	97.6	1.7	96.5	-1.9	98.3	1.0
guo	99.3	-0.7	98.4	-2.8	99.5	-0.4	99.5	0.3	99.9	-0.1
gup	99.5	-0.9	99.5	-1.0	98.6	-2.8	99.2	-1.6	97.3	-5.3
gur	97.3	2.4	98.3	-0.2	98.5	0.0	97.2	-0.8	98.8	0.8
guu	96.3	4.2	97.5	2.9	98.4	1.2	98.6	2.3	99.3	0.2
guw	97.9	1.8	99.1	0.4	97.7	3.7	98.7	1.8	99.3	1.0
gux	93.8	-0.2	96.4	2.5	93.4	-2.7	94.6	1.5	95.5	-0.3
guz	88.2	5.6	93.1	6.0	92.2	3.4	89.6	-0.4	92.9	2.8
gvc	98.4	0.1	99.6	-0.2	98.9	-0.3	98.4	-0.5	98.7	0.0
gvf	98.6	-2.7	99.7	-0.7	99.5	-0.9	98.9	-1.3	99.5	-0.5
gvl	97.2	-1.2	99.5	0.4	95.4	-4.7	95.8	-0.9	97.1	-3.2
gvn	98.4	-2.8	98.7	-2.5	99.2	-1.5	98.8	-0.6	99.5	-1.0
gvr	90.7	10.9	97.2	3.4	90.5	-9.0	86.7	-11.0	90.3	-10.8
gvs	90.3	3.0	96.1	5.3	93.6	1.8	91.2	-1.1	94.5	-0.8
gwi	99.7	0.6	99.7	0.1	99.4	0.5	99.3	1.3	99.3	0.3
gya	72.9	41.1	79.4	32.9	78.7	27.3	74.9	26.8	81.7	22.9
gym	99.0	0.7	98.9	1.3	98.5	0.6	98.2	1.6	99.4	0.7
gyr	98.6	-0.1	99.2	-1.4	99.3	-0.2	98.2	-1.9	99.4	-0.3
gyz	95.5	6.3	96.0	4.5	96.1	1.5	95.8	1.3	97.6	2.6
had	78.3	24.3	95.1	8.9	90.2	-0.2	90.3	4.3	91.3	-1.6
hae	82.6	4.0	93.6	-9.6	96.0	-2.2	90.6	-10.3	96.0	-1.7
hak	64.7	47.5	65.5	40.6	68.7	1.2	70.1	-28.3	81.9	-6.5
hap	64.4	42.6	69.1	46.1	64.8	16.6	52.8	8.9	66.7	24.1
hat	93.0	-9.7	96.6	-4.0	95.3	-5.2	94.1	-5.5	98.2	-0.4
hau	54.1	59.4	78.9	29.8	90.4	1.4	83.3	-11.4	93.4	3.1
haw	92.7	0.1	96.9	-2.0	95.5	-1.3	95.5	2.8	97.8	2.2
hay	77.5	-17.1	92.8	2.4	89.6	2.8	81.2	4.3	87.1	3.1
hbb	92.9	4.9	93.7	-1.5	93.3	3.3	92.3	6.0	94.6	1.4
hbn	97.2	-2.0	99.0	-0.4	98.0	-1.5	98.1	-0.1	98.2	-1.4
hbo	99.9	-0.2	100.0	-0.1	100.0	0.0	100.0	-0.0	100.0	0.1
hch	99.6	-0.0	97.9	-4.2	98.9	0.7	99.1	1.0	99.1	0.1
hdn	98.2	2.2	98.4	1.4	97.8	2.8	98.2	3.2	94.3	10.4
hdy	89.4	6.6	94.3	6.4	92.0	1.8	87.6	-0.4	92.3	-0.5
heb	99.1	-1.0	98.1	1.2	99.8	0.0	99.5	-0.4	99.8	-0.2
heg	84.3	-4.8	97.1	0.6	91.7	10.6	86.7	16.4	92.6	10.3
heh	76.1	6.5	94.1	3.7	87.1	-5.6	80.4	-5.2	88.4	-6.2
her	88.2	10.9	92.4	8.3	88.7	3.8	87.8	-1.7	91.5	4.3
hif	70.0	-21.6	82.8	-14.3	86.9	-4.1	85.5	4.6	93.4	2.3
hig	99.0	-0.9	99.6	-0.8	99.6	-0.7	99.6	-0.7	99.8	-0.1
hil	62.4	11.7	73.5	36.1	78.5	17.8	70.9	-1.7	83.0	13.7
hin	22.1	16.3	21.3	67.6	70.5	-5.2	51.8	-24.8	78.9	2.2
hix	99.2	-1.3	98.8	-2.4	99.3	-0.9	99.8	-0.0	99.5	-1.0
hla	95.4	2.3	98.7	0.5	96.6	2.1	95.3	3.4	97.3	1.8
hlb	83.4	-15.0	96.1	-7.0	96.7	-3.4	95.1	-1.8	96.1	-6.0
hlt	89.9	7.4	93.8	7.4	91.4	13.4	88.5	15.9	92.7	10.5
hmo	87.0	12.0	96.8	3.3	94.8	-3.7	93.0	-8.1	98.2	1.6
hmt	99.7	-0.5	99.5	-1.1	100.0	-0.0	100.0	-0.0	100.0	-0.0
hne	76.7	9.7	93.5	10.4	91.1	-3.6	86.4	-8.0	92.0	-5.9
hnj	93.7	0.6	96.0	0.8	95.4	1.7	90.8	-3.9	95.7	0.1
hnn	96.6	-3.2	99.0	-1.6	98.7	-1.6	98.6	-2.1	99.3	-1.0
hns	96.1	-2.3	97.7	-1.2	97.9	1.4	97.3	0.7	97.9	1.7
hoc	71.1	43.0	99.3	-0.9	99.6	0.5	98.5	2.4	99.6	0.4
hop	99.7	-0.5	99.3	-1.4	99.5	-0.8	99.9	0.0	99.9	-0.1
hot	98.3	-1.1	99.5	0.1	98.5	-0.9	98.3	-0.7	99.3	0.5
hoy	88.9	-0.6	97.9	-2.7	99.1	0.0	97.8	1.9	99.4	0.4
hra	78.3	16.5	89.2	12.5	86.7	12.4	75.1	16.3	89.6	10.1
hrv	37.7	-3.6	43.9	6.7	49.4	-1.8	38.7	-14.8	59.8	-1.7
hrx	95.2	-0.9	97.6	1.3	93.1	-3.3	94.5	0.3	95.3	-2.7
hsb	58.4	26.1	66.8	-10.5	88.3	4.4	71.0	27.9	91.5	10.8
hto	97.5	-1.6	98.7	-1.7	99.2	0.3	99.2	0.1	99.6	-0.0
hub	83.9	-3.3	95.4	-4.4	93.0	-1.2	87.5	-2.4	94.1	1.0
hui	98.4	-2.6	98.6	-2.1	99.3	-0.7	99.2	1.0	99.8	0.3
hun	92.7	5.5	93.7	5.5	95.8	-0.3	95.4	2.2	98.4	1.2
hus	99.6	0.2	99.7	-0.3	99.2	-0.3	99.4	-0.2	99.7	-0.1
huu	97.1	1.9	98.6	-0.6	98.4	-1.1	98.6	0.3	99.0	-0.1
huv	98.7	0.1	99.9	-0.3	99.0	-0.3	99.2	-0.1	99.5	-0.5
hvn	81.1	6.7	86.7	15.9	85.1	15.0	74.9	-8.1	83.8	7.8
hwc	91.8	0.9	97.4	-1.4	94.9	3.4	91.8	1.1	97.6	-1.0
hye	90.5	-7.8	74.0	5.1	94.4	-6.2	89.7	-10.2	96.9	-1.5
hyw	85.9	-11.4	62.8	-43.7	96.6	0.2	88.1	10.1	97.6	2.0
iai	92.9	6.5	97.8	0.5	96.3	2.6	95.4	4.2	96.4	2.3
ian	99.3	-0.3	99.6	-0.2	99.5	0.4	99.2	0.0	99.0	-0.4
iba	77.7	-3.4	86.4	-8.1	83.4	-9.2	79.3	-9.2	83.9	2.5
ibb	97.6	3.4	97.6	1.4	94.7	-0.0	96.0	0.1	96.8	-0.5
ibo	92.7	4.9	98.2	0.4	96.0	-1.0	94.4	1.5	98.0	-0.1
icr	90.2	7.4	97.7	-3.6	96.9	-1.7	93.2	-3.9	98.7	-0.4
idd	99.1	-1.8	99.7	-0.6	100.0	0.0	100.0	0.1	99.9	-0.3
ido	47.6	21.5	70.4	24.8	87.0	4.5	79.2	6.4	93.9	1.7
idu	95.3	6.4	97.1	4.5	97.8	1.3	97.1	-1.4	98.6	-0.2
ifa	93.7	-5.9	97.4	0.4	97.9	2.1	98.0	2.2	99.2	1.5
ife	97.1	1.8	97.3	-3.7	97.0	-3.5	97.9	-1.6	97.6	-2.3
ifu	97.5	-3.6	99.2	-0.8	99.5	-0.3	99.5	0.2	99.8	0.1
ify	93.3	-1.5	96.9	3.8	96.0	4.8	96.3	6.3	96.8	5.1
igb	94.8	3.3	95.6	6.0	97.5	0.5	95.0	-3.1	98.5	-0.7
ige	96.4	3.5	97.5	3.0	96.9	2.6	96.2	2.8	97.9	2.6
igl	98.4	2.0	98.6	2.0	97.2	1.2	98.9	1.6	99.1	0.9
ign	99.1	-0.8	99.2	-0.9	99.2	-0.2	99.3	0.8	99.4	-0.5
iii	100.0	0.0	99.8	0.3	99.6	-0.9	100.0	0.0	99.6	-0.9
ijn	89.0	6.7	93.7	7.3	93.8	2.8	92.3	2.1	96.2	2.4
ike	61.9	40.7	66.1	47.2	96.9	-0.5	97.2	3.6	96.9	2.5
ikk	96.1	1.5	98.4	2.3	97.3	2.2	95.8	4.2	98.6	2.1
ikt	86.3	-22.0	96.7	-3.9	96.1	-5.2	95.2	-6.6	96.9	-4.1
ikw	98.0	2.2	97.6	-2.4	97.4	0.8	96.2	1.3	98.1	1.4
ilb	58.3	5.3	79.3	11.4	81.1	-0.6	71.4	-7.4	83.0	-5.2
ile	60.9	-14.5	82.7	13.4	83.8	-7.4	78.8	-3.7	88.9	-9.2
ilo	73.5	-1.7	85.5	9.3	84.4	4.2	81.0	6.1	89.1	9.4
imo	96.0	-3.6	98.5	-2.4	98.0	-1.5	96.7	1.5	98.4	-1.5
ina	69.5	-36.9	89.7	-11.9	87.4	-15.1	86.9	-9.4	95.0	-5.8
inb	97.8	3.1	99.0	-1.1	98.3	0.7	97.1	4.7	98.1	2.7
ind	49.2	-30.6	66.4	-28.9	74.4	-18.3	61.0	-35.1	86.4	-8.3
inh	80.8	-5.9	70.8	37.4	86.3	-11.9	74.5	-7.0	84.2	-14.9
ino	99.4	-1.1	99.5	-1.1	99.8	-0.1	99.7	0.1	99.8	-0.1
iou	99.5	-1.0	99.8	-0.4	99.6	-0.7	99.9	0.1	99.7	-0.4
ipi	98.5	-2.2	99.3	-1.4	99.5	-0.9	99.4	-0.4	99.7	-0.1
iqw	72.8	-0.7	90.1	-0.2	87.6	-6.5	79.4	-3.2	91.7	-7.2
iri	91.7	-15.0	92.0	-14.8	96.2	-7.3	99.0	-1.6	99.3	-0.9
irk	89.8	5.9	92.9	2.2	93.8	7.0	92.4	-0.5	90.9	-4.1
iry	96.4	-2.7	99.5	-0.5	98.9	-0.3	98.9	-0.2	99.6	-0.2
isd	96.3	-1.9	99.2	0.2	98.3	0.5	98.5	1.5	99.3	0.2
isl	88.6	-12.0	95.5	-2.7	93.9	-2.0	90.5	-3.1	95.8	-1.0
isn	90.2	-12.3	97.7	-4.0	97.4	-2.1	93.3	-8.7	97.9	-2.1
ita	77.0	-22.8	85.3	-22.5	87.9	-15.2	74.2	-32.1	87.4	-18.9
itv	96.2	-6.2	99.3	-1.0	99.3	-1.0	99.1	-0.7	99.7	-0.2
ium	93.0	10.0	95.2	3.0	94.8	-3.3	96.1	1.2	98.4	1.6
ivb	91.8	1.1	97.2	2.0	97.1	0.2	95.5	0.1	97.4	3.1
ivv	91.8	-9.2	98.4	-2.2	98.7	-1.2	97.6	0.0	98.3	-2.2
iws	99.7	-0.6	99.9	-0.2	99.9	-0.1	99.9	-0.0	99.9	-0.1
ixl	99.1	0.7	99.5	-0.7	98.9	0.6	99.0	0.8	99.6	0.5
izh	83.2	-0.4	94.9	1.2	91.1	-5.2	77.9	-20.6	83.6	-23.1
izr	92.8	8.4	95.8	4.9	94.7	3.8	92.7	0.2	95.2	1.1
izz	68.5	10.7	89.4	10.1	86.7	10.0	75.5	15.4	90.2	11.2
jaa	98.6	-2.4	99.2	-1.2	99.1	-1.4	99.7	0.1	99.6	-0.5
jab	95.9	5.8	98.6	2.3	97.1	3.3	95.3	5.3	97.8	2.4
jac	99.1	0.9	99.8	-0.1	99.4	0.6	99.3	0.1	99.6	0.3
jae	97.6	4.1	98.6	1.9	97.0	3.3	97.2	3.5	98.2	2.6
jam	84.4	-11.7	95.2	0.8	91.8	-3.0	85.8	-7.6	95.1	-0.7
jav	70.5	-5.5	89.7	3.7	86.3	-2.8	76.1	14.9	92.7	4.1
jbn	99.3	-1.0	96.3	-7.1	99.6	-0.5	98.6	-2.4	99.2	-1.3
jbo	90.2	14.0	92.2	3.1	94.6	1.5	95.1	-0.7	96.1	-2.8
jbu	97.9	-3.8	99.3	-1.4	99.4	-1.2	99.3	-1.0	99.7	-0.6
jen	99.5	-0.7	99.6	-0.6	99.8	-0.1	99.8	-0.0	99.9	0.1
jgk	96.6	0.8	98.7	1.0	97.0	-0.3	95.8	0.4	97.4	-1.5
jgo	98.7	2.1	99.3	-0.4	98.7	0.8	98.5	1.6	99.1	0.7
jic	96.5	2.8	98.7	-0.0	96.6	3.9	96.2	4.0	98.2	2.5
jid	95.5	1.3	99.2	-0.1	95.7	3.8	93.4	5.9	97.2	2.9
jit	90.5	3.7	94.2	2.3	90.7	-5.6	91.1	-1.3	87.2	-16.0
jiv	95.2	-0.3	98.3	-1.9	97.2	-1.2	97.4	0.6	98.4	-0.4
jmc	78.3	14.0	91.4	10.5	89.2	8.9	86.1	3.4	89.3	8.8
jni	92.9	-11.1	99.2	-1.0	96.8	-2.5	93.9	-7.1	97.7	-2.4
jpn	99.0	0.9	95.7	-4.4	94.6	-3.2	99.1	0.7	98.6	-0.9
jra	31.6	71.5	61.4	54.4	94.1	5.1	94.4	-0.1	96.8	2.0
jun	95.2	-5.8	99.8	-0.1	99.8	-0.3	99.7	0.2	99.9	0.1
juy	98.4	-0.4	99.8	-0.1	99.8	0.0	99.2	-1.1	99.7	0.4
jvn	92.6	1.1	97.9	0.1	95.5	0.4	91.3	-3.3	97.0	-0.8
kaa	37.3	55.4	73.6	31.0	89.1	-1.5	88.3	-1.6	93.8	-1.5
kab	72.8	-36.0	89.8	-12.1	91.2	-4.0	90.8	3.1	94.0	5.1
kac	91.2	10.5	91.6	14.7	93.6	6.1	92.6	3.4	94.5	2.7
kaj	96.2	-0.4	99.0	0.0	97.9	0.5	96.9	0.6	98.4	1.3
kak	97.7	-2.0	98.9	-1.9	99.3	-0.9	99.2	-0.5	99.2	-0.6
kal	48.4	4.5	65.2	41.7	55.4	-0.1	46.7	15.9	69.2	-0.5
kam	92.9	9.4	95.3	5.4	93.7	2.6	93.1	-1.1	94.9	2.4
kan	80.1	-17.5	81.6	-24.8	94.7	-2.8	91.3	-6.8	97.3	0.2
kao	98.1	-3.0	98.2	-3.5	99.2	-0.9	99.3	0.5	99.5	-0.2
kap	94.7	1.0	91.3	14.7	97.0	1.2	95.3	3.8	98.0	0.4
kaq	98.0	-1.4	99.3	-1.0	98.5	-0.2	98.4	2.1	99.2	0.8
kat	87.8	-6.4	86.8	-5.9	94.4	-5.7	91.8	-13.2	98.0	-2.7
kaz	83.4	-11.2	84.7	-15.1	92.8	1.2	90.9	0.6	95.1	0.8
kbc	99.0	1.9	99.1	-0.3	98.7	1.4	98.1	3.4	98.8	2.0
kbd	84.4	10.9	87.5	-0.0	89.3	11.2	77.2	28.0	91.0	9.0
kbh	99.5	0.3	99.7	-0.5	98.7	-1.8	99.5	-0.0	99.5	-0.7
kbm	98.6	-2.0	99.6	-0.4	99.1	-0.7	99.4	0.5	99.6	-0.3
kbn	98.6	-1.2	99.4	-1.1	99.3	-0.6	99.0	-0.1	99.2	-1.1
kbp	96.3	-2.8	96.9	-3.1	90.9	-11.4	87.5	-17.5	97.8	-0.2
kbq	99.0	-1.8	98.6	-2.7	99.1	-1.6	99.4	-0.4	99.4	-0.7
kbr	92.2	4.7	94.2	4.3	92.8	-0.8	89.1	-6.1	91.9	-3.4
kca	99.2	-1.0	99.4	-0.8	99.7	-0.1	99.8	0.1	99.9	0.2
kcg	92.9	7.9	95.7	5.9	92.2	-4.7	86.7	-14.0	92.9	-7.8
kck	78.0	16.9	92.4	12.1	90.4	5.0	84.3	8.3	90.8	8.2

Table 6: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
kdc	41.3	-6.8	75.4	-15.8	74.4	-10.5	56.0	-12.3	71.8	-21.3
kde	91.2	3.6	98.1	-2.1	96.1	-1.5	93.5	-4.9	97.3	-1.6
kdh	95.7	5.2	96.7	1.7	96.0	4.8	95.2	5.0	95.6	6.2
kdi	86.5	11.4	91.5	11.3	91.5	5.2	85.2	-2.0	92.1	4.0
kdj	93.3	1.3	96.0	-4.0	95.6	0.1	94.4	-1.8	94.7	-1.8
kdl	89.0	-6.4	98.2	-1.8	96.1	-1.0	91.2	-3.4	97.3	-0.7
kdn	55.0	33.3	84.6	19.8	78.0	25.7	69.1	28.1	80.4	27.2
kdt	88.5	8.9	88.1	16.7	99.2	0.0	98.4	-0.8	98.8	-0.5
ked	79.4	-0.2	93.3	0.8	91.4	-2.4	83.9	-13.3	90.5	-5.7
kek	98.4	-2.1	99.5	0.1	99.6	0.4	99.4	0.4	99.8	0.2
ken	99.3	-1.1	99.0	-2.0	99.5	-0.1	99.3	0.6	99.5	0.0
keo	98.5	-1.4	98.9	-2.2	99.0	-1.7	98.6	0.7	99.8	-0.2
ker	96.4	1.9	96.7	-1.9	96.1	0.7	94.3	-1.4	95.7	-0.6
kew	83.5	-3.0	85.7	19.9	90.1	9.3	86.1	2.8	93.1	5.6
kez	96.1	-7.0	98.6	-2.8	99.4	-1.2	99.9	0.0	99.8	-0.2
kfa	74.1	28.6	86.5	9.8	60.9	54.7	71.8	40.9	96.6	4.0
kfb	92.2	-10.1	99.2	-1.5	99.7	-0.2	99.1	0.6	99.9	0.2
kff	90.9	5.7	96.0	4.9	98.1	0.6	96.2	-3.5	98.0	0.8
kfi	94.8	2.0	95.8	3.8	99.0	0.8	98.6	1.5	98.9	0.9
kfp	81.4	1.5	92.0	4.3	95.4	0.6	93.7	1.7	94.7	-0.4
kfs	63.5	1.0	95.4	-4.1	96.4	2.4	88.7	3.1	97.8	1.1
kfw	95.2	-0.6	98.0	-3.6	97.1	-1.5	95.4	0.4	97.2	-1.3
kfx	82.1	6.1	96.2	3.0	96.0	2.4	92.6	2.4	97.6	2.5
kfy	69.8	-12.3	91.6	4.2	85.8	-8.9	87.0	-1.3	92.0	-3.3
kgf	99.5	-0.7	99.6	-0.5	99.9	0.2	99.8	0.1	99.8	-0.1
kgk	95.4	-2.9	98.8	-1.6	98.1	1.1	95.5	5.3	97.8	2.3
kgp	99.0	-1.5	98.5	-2.8	99.6	0.8	99.8	0.4	99.9	0.1
kgr	91.9	-1.3	97.5	3.3	95.2	3.9	94.7	3.0	95.3	-0.5
kha	92.0	10.5	94.1	8.6	93.0	3.0	92.6	-0.2	94.3	1.9
khg	38.8	6.6	42.1	-12.3	54.8	-5.0	52.3	-8.2	39.4	12.6
khk	86.4	9.7	81.4	29.7	94.3	0.5	91.7	-5.5	96.5	-0.3
khm	88.5	-13.1	86.6	-21.4	99.2	-0.1	98.7	1.3	99.5	0.6
khn	94.6	-7.6	99.1	-1.8	99.6	-0.2	99.4	-0.1	99.7	-0.3
khq	94.6	-8.9	99.2	-1.2	99.1	-0.6	99.0	-0.3	99.4	-1.0
khs	94.6	-8.9	98.4	-3.0	99.1	-0.5	99.1	1.1	99.3	0.5
kht	100.0	0.0	99.9	-0.2	99.8	-0.3	100.0	0.0	100.0	0.0
khw	76.1	-11.2	82.3	-4.5	89.6	0.6	76.2	-18.3	90.0	-8.8
khy	96.2	1.2	98.1	-3.1	95.8	-6.3	94.5	-5.8	95.2	-6.5
khz	95.3	-1.7	98.4	-2.0	96.2	-3.8	94.4	-2.5	95.5	-5.5
kij	88.8	1.5	93.2	3.0	92.9	7.1	91.8	10.9	93.4	7.2
kik	96.6	1.8	96.8	-0.2	96.1	-2.7	96.6	-0.6	97.8	-1.3
kin	39.8	23.3	61.6	33.3	57.1	9.9	49.3	-7.2	63.7	17.0
kir	73.5	-13.5	86.4	-7.5	93.4	3.3	87.2	5.6	95.8	0.5
kiu	68.3	-26.3	73.2	-27.2	74.5	-15.0	61.0	-35.0	64.7	-47.2
kix	95.3	6.7	96.7	2.9	93.5	-1.9	93.2	3.3	96.4	2.1
kiz	95.5	2.4	98.5	-1.8	97.7	0.1	97.0	0.7	98.7	-0.0
kjb	97.1	-3.7	99.4	-1.0	99.1	-1.1	98.8	-1.4	99.4	-0.8
kje	94.4	-1.4	98.4	0.2	97.1	1.7	95.0	2.8	97.4	1.9
kjh	89.0	2.5	94.8	7.4	94.6	-6.8	91.4	-7.5	96.0	-4.8
kji	92.2	4.0	97.2	1.3	94.5	-4.5	92.6	-4.1	93.8	-8.7
kjo	93.2	0.6	98.3	0.1	98.4	0.7	97.9	-0.3	98.5	0.4
kjs	83.7	-2.4	88.5	-18.3	91.8	-10.8	87.5	-5.1	94.0	-6.5
kkc	94.5	-10.1	96.0	-7.4	96.5	-6.1	99.5	-0.2	99.7	-0.5
kki	72.2	15.0	92.8	8.0	85.0	-4.8	77.3	-5.6	85.9	1.5
kkl	94.1	-4.1	98.4	-1.3	97.6	-1.2	96.0	0.0	96.6	-2.7
kle	94.9	-1.7	99.1	-1.7	99.7	-0.0	98.6	2.1	99.7	0.3
klt	98.8	-1.9	99.2	-1.6	99.5	-0.2	99.4	-0.5	99.5	-0.5
klu	96.1	6.0	97.2	3.7	96.5	-1.5	96.3	-1.6	96.9	-2.9
klv	98.7	1.1	98.8	-0.8	97.9	1.0	98.4	1.1	98.7	1.7
kmb	94.9	2.3	97.1	1.4	96.4	-4.1	96.8	-2.7	97.2	-0.4
kmc	97.5	-4.1	97.6	-4.0	99.1	1.0	98.6	-0.0	99.1	0.1
kmd	97.4	-4.7	99.4	-1.1	99.7	-0.2	99.8	0.2	99.9	0.0
kmg	97.7	3.5	98.4	2.2	98.0	1.9	97.9	2.7	98.6	1.1
kmh	99.4	0.2	99.7	-0.1	99.2	0.8	99.2	1.3	99.4	1.0
kmk	95.9	-4.3	98.6	-2.3	98.5	-1.4	98.4	-0.7	99.2	-0.3
kml	95.3	7.5	95.5	4.7	92.1	-2.5	95.7	2.2	97.1	-0.7
kmm	87.4	9.3	90.9	8.6	90.4	8.7	84.7	5.0	92.6	7.0
kmo	97.6	-1.5	99.3	-1.0	98.1	1.3	97.2	1.1	98.6	1.4
kmr	94.5	3.1	97.2	0.8	93.1	4.4	94.0	0.3	98.4	1.0
kms	98.8	-0.9	99.6	-0.5	99.2	-0.1	98.9	0.0	99.6	0.3
kmu	98.2	-2.4	99.3	-1.3	98.4	-1.2	98.9	0.0	99.1	0.7
kmw	93.4	5.9	89.5	6.1	96.4	-0.1	95.5	1.1	96.5	-0.2
kne	84.1	-5.5	96.1	2.7	91.5	8.0	89.0	8.9	95.3	4.9
knf	97.4	-0.2	98.9	-0.3	97.5	1.0	97.7	1.1	98.6	2.1
knj	98.7	-0.8	98.9	-0.8	98.9	0.3	98.0	-1.0	99.2	0.3
knk	86.1	-1.3	95.3	2.3	91.5	1.0	86.5	-6.7	93.0	1.3
kno	96.8	0.9	98.9	1.1	97.6	-0.2	97.5	0.9	98.2	-0.4
kog	98.5	-2.6	99.2	-1.7	99.4	-1.1	99.6	-0.4	99.8	-0.1
koi	54.8	29.5	52.0	55.3	80.0	6.5	57.3	39.8	86.9	7.1
kor	99.4	-1.1	82.7	-23.0	97.8	-2.1	99.2	0.7	99.1	-1.2
kos	90.4	3.2	95.2	6.5	92.3	4.7	91.4	10.6	96.7	3.7
kpf	98.8	-0.8	99.8	-0.5	99.3	0.6	99.1	1.0	99.3	0.5
kpg	93.6	10.7	94.2	4.8	93.9	5.6	94.7	9.9	95.0	7.9
kpj	99.7	-0.5	99.4	-1.0	98.7	-2.5	99.1	-1.6	98.9	-2.1
kpm	98.0	-0.3	98.2	0.0	98.3	1.0	98.4	-1.1	99.0	-0.7
kpr	95.4	-5.1	98.5	-2.3	98.6	-0.5	96.5	-2.9	99.2	-0.1
kpv	92.3	-13.3	93.1	-12.6	98.7	-1.5	92.6	-12.4	98.3	-2.1
kpw	98.1	1.9	98.2	-0.9	97.9	2.9	96.9	3.0	98.1	2.3
kpx	96.8	-3.6	97.9	-2.5	98.2	-0.6	98.0	0.9	98.9	0.9
kpz	98.4	-2.7	97.6	-4.5	99.4	-0.8	99.8	0.1	99.7	-0.2
kqc	97.0	-1.6	98.5	-1.8	98.2	-0.5	97.2	-2.1	98.8	0.4
kqe	86.2	-9.9	98.5	0.3	97.0	0.3	93.7	2.1	97.9	-1.5
kqf	94.9	0.0	98.8	0.0	96.8	-0.5	95.6	0.0	97.5	0.5
kqn	71.7	8.6	80.9	9.5	79.8	-7.0	72.9	-11.7	86.6	-0.5
kqo	98.1	2.0	96.8	-0.5	96.7	0.4	95.9	-1.7	97.3	-0.5
kqr	93.4	-7.7	98.2	-3.0	98.8	-0.8	98.4	0.3	99.4	-0.5
kqs	90.0	7.6	94.7	7.7	90.5	7.4	90.9	5.2	92.1	0.6
kqw	89.8	-9.1	98.6	0.6	96.9	0.7	94.9	-0.6	97.4	0.0
kqy	89.5	-5.1	95.6	1.0	94.8	-1.5	91.3	0.2	94.9	1.5
krc	66.1	-5.7	79.9	-10.0	91.9	2.8	84.4	9.8	94.8	3.6
kri	87.5	-5.6	95.6	3.7	94.9	0.5	88.1	-11.0	95.3	-0.0
krl	97.8	-0.4	98.0	-3.6	97.8	-1.7	98.0	-0.1	99.3	-0.1
krr	85.2	-16.5	85.1	-20.4	98.4	-0.6	97.7	0.9	98.7	0.1
krs	96.9	5.4	98.0	2.9	97.6	1.5	97.5	2.6	98.0	3.0
kru	72.4	13.9	85.6	0.0	90.4	4.4	85.7	1.5	92.4	3.6
krw	95.2	4.5	96.2	3.2	97.4	2.0	95.6	0.0	97.6	2.3
ksb	82.0	1.9	93.4	-0.3	88.9	-6.8	85.2	-3.0	92.3	1.4
ksc	90.7	-15.4	99.0	-1.8	98.7	-2.1	99.1	-1.1	99.4	-0.5
ksd	88.7	-3.7	97.0	2.5	94.5	5.7	92.8	5.7	95.4	7.6
ksf	98.0	2.0	98.3	2.3	97.5	-0.2	97.8	0.8	97.9	1.6
ksh	60.6	-2.9	63.9	-15.9	78.1	12.3	61.3	37.0	82.3	14.6
ksj	99.5	-0.2	99.8	-0.5	99.1	0.1	99.5	0.1	99.6	-0.4
ksr	98.2	-2.8	99.0	-1.1	99.2	-1.0	99.3	0.6	99.7	-0.2
kss	99.0	-1.7	98.1	-3.8	99.5	-1.0	99.9	-0.2	99.9	-0.3
ksw	97.9	1.8	98.4	2.9	99.3	-1.4	99.9	-0.2	99.9	-0.2
ksz	71.6	30.4	73.5	-26.7	86.7	-13.6	84.8	-13.8	88.3	-11.7
ktb	84.7	4.9	90.2	1.8	98.9	-0.6	94.7	2.4	98.3	0.2
ktj	95.8	-2.0	98.8	-1.2	97.3	-2.4	96.3	0.4	97.9	1.5
ktm	95.4	-1.8	99.6	0.1	96.8	-0.5	95.2	3.6	97.0	-1.5
kto	96.8	-4.9	99.3	-1.3	99.4	-0.7	99.1	0.1	99.2	-1.2
ktu	53.2	5.9	74.1	-24.2	73.9	-30.8	74.4	-29.5	77.7	-32.1
kua	88.5	1.6	96.1	2.5	92.8	5.6	88.4	4.2	94.9	4.6
kub	88.0	-2.4	95.1	7.0	94.0	5.0	91.3	-0.4	93.9	2.9
kud	96.7	-5.2	99.2	-1.1	97.6	-3.5	97.9	-0.3	98.4	-2.1
kue	94.2	5.5	98.8	0.3	97.4	1.6	94.9	0.2	96.8	-0.3
kuj	89.6	4.3	93.7	5.1	93.0	1.4	91.2	-4.3	93.6	2.1
kum	69.4	-18.9	71.1	-34.1	91.2	-1.3	80.4	-3.7	92.6	-1.9
kup	96.4	-6.5	99.0	-1.6	99.1	-0.9	99.4	0.3	99.5	0.4
kvg	98.2	1.5	98.7	-2.1	97.2	2.8	97.5	4.6	97.5	4.0
kvj	96.8	3.0	95.8	-2.7	97.5	1.5	97.7	1.0	98.8	0.6
kvn	94.5	0.2	96.3	1.3	97.7	0.3	94.9	-2.7	97.9	0.0
kvq	98.5	2.9	97.3	-1.8	99.1	0.5	99.3	-0.2	99.5	-0.1
kvy	91.3	16.0	91.3	16.0	91.2	15.8	90.8	16.4	83.0	13.9
kwd	97.9	-0.3	99.3	-0.8	98.4	0.4	97.8	-0.7	98.7	0.0
kwf	98.8	-0.0	97.2	2.4	98.7	0.4	99.6	0.7	99.5	0.4
kwi	96.6	-2.8	97.1	-3.2	98.4	1.2	98.2	1.7	98.2	1.0
kwj	97.5	-3.5	99.4	-1.1	99.5	-0.3	98.9	-0.2	99.5	-0.5
kwk	98.6	-2.0	99.0	-1.5	97.9	-3.5	99.7	0.1	99.3	-1.1
kxc	98.6	-2.7	99.4	-1.3	99.6	-0.5	99.4	-0.4	99.8	-0.2
kxf	95.9	7.6	97.4	2.5	97.5	3.9	99.2	1.0	99.2	0.1
kxm	96.9	-5.0	99.0	-1.1	100.0	0.0	99.9	0.1	100.0	0.0
kxv	97.7	-0.3	99.4	0.2	99.7	0.0	99.6	0.3	99.9	-0.0
kxz	98.8	-0.3	98.8	-0.7	98.3	-2.1	99.0	-0.1	98.4	-2.2
kyc	96.8	-3.2	97.5	-4.7	97.7	2.3	95.5	6.2	97.9	2.9
kyf	97.0	1.0	98.4	-0.6	97.7	2.0	96.4	1.5	98.8	1.2
kyg	99.7	-0.4	99.8	-0.5	99.7	-0.0	99.3	1.2	99.6	0.7
kyj	83.5	5.3	92.9	9.3	85.2	-2.7	82.8	-1.6	88.0	-5.0
kyq	99.0	0.6	98.5	-2.1	98.1	0.8	97.4	3.4	99.0	1.1
kyu	98.9	1.2	97.1	-2.5	98.7	2.1	98.7	2.3	98.6	2.0
kyv	73.3	11.7	95.8	-2.0	97.7	2.4	94.9	2.8	98.2	3.2
kyz	99.4	-0.4	98.8	-2.2	99.2	1.0	98.6	2.0	99.2	1.5
kze	97.5	-2.1	98.9	-1.9	98.6	0.2	97.8	1.0	98.6	0.1
kzf	80.4	10.7	90.9	11.2	81.1	-6.9	82.2	3.5	86.1	-3.0
kzr	91.9	2.3	97.2	4.5	92.7	-3.6	92.3	-2.2	92.0	-7.1
lac	94.7	-8.9	97.8	-3.9	99.5	-0.3	99.4	0.1	99.6	0.0
lad	83.0	-16.1	91.1	-11.9	92.9	0.5	82.9	22.4	93.6	5.4
lai	75.0	6.9	82.7	-7.8	87.3	-6.3	77.2	-13.2	86.8	-16.2
laj	85.9	10.4	91.7	12.7	89.9	8.2	84.5	-0.4	91.7	4.0
lam	92.4	7.2	96.7	-1.9	96.8	-0.2	95.7	5.9	98.0	3.5
lao	94.6	9.4	94.5	10.4	92.0	3.2	94.6	7.4	94.8	2.9
lap	86.1	21.4	87.7	21.3	88.7	5.3	86.2	0.1	89.8	2.2
las	99.5	-0.8	99.5	-1.0	99.6	-0.8	99.2	-1.2	99.7	-0.6
lat	80.6	-15.2	87.5	-17.2	93.0	-4.8	92.4	-0.7	95.8	-4.4
law	97.7	-4.4	99.2	-1.7	99.8	-0.5	99.8	-0.2	99.8	-0.3
lbb	91.4	-4.0	98.3	0.6	96.4	0.9	94.1	3.4	97.2	0.4
lbf	91.7	3.9	97.8	-2.1	98.2	0.6	97.8	2.1	99.2	0.9
lbk	89.4	-5.2	96.7	3.6	96.1	2.2	92.5	4.8	95.5	1.5
lbm	86.1	2.1	97.1	-4.4	97.5	-1.8	96.4	-2.9	97.0	-3.9
lbr	78.1	-19.5	87.6	-8.8	93.8	0.0	93.9	-0.3	94.9	1.7
lbw	98.2	-1.4	99.2	-1.0	98.8	-0.9	98.5	0.1	99.6	0.3
lcm	87.8	-19.4	99.2	-0.7	97.6	-1.5	96.7	-2.1	98.8	-0.7
lcp	86.8	10.6	98.5	3.0	99.5	0.5	99.1	-0.1	99.5	0.4
leb	69.2	2.8	89.2	1.7	80.6	-13.8	71.8	-18.7	78.9	-23.2
lee	97.9	1.6	99.2	1.2	98.2	-1.0	98.1	-1.6	99.1	0.2
lef	97.5	-0.9	97.8	-0.1	97.5	1.1	98.5	-1.1	99.2	-1.3
leh	73.9	6.6	90.2	10.0	87.1	2.3	80.5	15.9	90.0	8.5
lem	89.4	8.6	94.9	6.7	92.1	-2.4	94.6	3.4	95.0	-0.5
leu	97.3	-2.1	99.4	-0.8	97.8	0.8	97.5	3.2	99.0	0.6
lew	82.1	-2.4	92.5	2.6	87.4	-2.8	82.8	-0.6	87.6	-6.8
lex	93.2	-4.2	98.1	-0.2	97.5	0.4	97.0	0.3	98.3	1.3
lez	66.9	-7.7	76.7	8.8	81.0	8.2	72.9	9.0	85.0	8.2
lfn	64.0	-20.2	82.2	9.0	79.2	5.3	72.8	-4.2	84.3	12.7
lgg	70.1	-3.7	68.5	-29.4	81.0	-23.8	83.3	-11.1	81.8	-26.6
lgl	96.4	-6.4	98.8	-1.8	98.4	-1.9	97.5	-2.2	99.4	-0.8
lgm	89.5	-1.5	95.1	0.1	90.2	-11.3	90.7	-8.4	94.4	-3.2
lhi	96.3	4.8	97.3	5.1	96.8	4.2	95.8	4.6	97.0	2.5
lhm	85.2	17.9	88.0	21.4	88.1	20.7	82.4	13.0	87.8	17.5
lhu	99.6	0.3	99.5	-0.6	99.6	-0.1	99.8	0.3	100.0	0.0
lia	85.7	13.8	87.4	10.1	92.0	-0.7	90.9	-2.6	94.2	-2.4
lid	96.5	-2.0	99.5	0.1	97.9	2.0	96.3	4.9	97.4	3.8
lif	66.7	49.8	66.4	49.0	99.8	-0.1	99.6	-0.4	99.8	0.4
lij	87.7	3.3	92.4	-6.1	91.1	-2.9	87.5	13.7	96.6	0.4
lim	50.8	6.3	58.7	-1.0	55.6	34.2	16.1	60.9	58.9	54.3
lin	77.5	-10.0	92.3	-1.3	89.5	-4.0	87.1	-1.7	93.8	0.3
lis	51.0	-63.7	40.5	-74.5	95.1	-3.9	97.7	-3.0	98.9	-1.9
lit	82.8	-3.6	72.0	-37.5	88.0	-6.9	81.3	-17.4	94.6	-2.2
liv	89.4	13.6	83.7	-8.1	83.9	-8.2	81.5	-12.1	81.2	-14.8
lje	86.7	11.0	91.0	13.4	87.8	12.2	84.5	7.9	88.6	13.2
ljp	45.6	-39.8	65.8	-20.9	79.4	6.0	75.2	2.9	70.9	-16.0
lki	37.5	4.9	46.2	9.9	72.0	3.7	45.6	-3.6	70.5	-22.9
lkt	93.1	-0.3	94.7	1.8	94.6	0.8	96.5	4.3	96.2	3.1
llb	72.0	22.6	89.7	14.8	84.4	7.5	76.2	7.9	86.6	10.8
lld	79.6	7.9	87.9	5.7	84.7	-10.3	70.0	-20.9	83.3	-19.1
lle	95.0	-7.0	98.2	-2.7	97.8	-3.2	97.3	-3.1	98.6	-1.2
llg	63.4	15.6	85.8	-0.1	80.4	2.1	69.9	20.1	84.6	13.1
lln	99.5	-0.8	99.7	-0.5	99.5	-0.6	99.6	0.2	99.8	-0.2
llp	92.5	4.0	96.9	0.3	94.7	7.9	94.2	9.9	95.6	6.2
lme	87.1	21.6	89.1	19.5	87.2	11.8	84.7	10.3	86.9	12.4
lmk	92.6	8.6	96.5	6.5	94.9	1.5	92.0	1.9	95.9	-1.8
lml	92.7	7.3	96.6	4.5	92.5	2.7	90.2	3.2	92.4	-3.5
lmo	73.7	-13.5	84.3	-15.0	85.9	-1.2	71.5	14.6	92.2	4.9
lmp	98.7	1.4	99.7	0.1	99.0	0.8	98.5	1.4	99.0	-0.3
lnd	93.5	6.9	94.5	5.1	95.0	6.3	94.1	0.4	96.4	1.1
lnl	93.2	9.8	94.8	9.4	93.4	-1.5	94.5	-3.2	94.7	-2.9
lns	99.7	-0.7	99.9	-0.2	100.0	0.0	99.9	0.1	100.0	-0.1
lob	88.3	11.1	90.4	3.4	92.5	0.8	91.8	0.8	94.4	1.5
loe	90.2	6.9	93.2	9.7	84.4	-7.3	83.6	-9.8	86.6	-5.0
lok	97.7	-3.9	99.3	-1.4	99.5	-0.7	99.6	-0.2	99.7	-0.2
lol	91.2	0.3	94.6	0.9	93.4	4.0	91.7	5.1	94.5	6.4
lom	97.2	0.3	99.0	0.1	96.0	0.5	96.0	1.4	95.4	-3.8
lon	80.8	10.9	91.4	10.0	87.9	6.3	83.8	6.8	88.9	4.3
lot	91.5	5.1	95.1	2.4	90.5	-4.8	89.7	-4.1	91.0	-7.8
loz	80.3	-4.1	91.6	11.4	87.9	9.2	83.4	-2.0	89.8	8.4
lrc	87.5	12.6	88.5	9.3	93.1	-3.7	91.0	7.6	94.1	3.2
lsi	96.7	-2.9	98.3	-0.1	96.7	-1.3	94.6	-6.2	97.5	-1.0
lsm	86.1	6.3	92.7	6.9	93.1	3.1	89.1	6.1	94.0	5.3
ltg	85.0	6.7	89.7	3.7	85.8	-4.0	83.7	-1.7	89.0	-5.6
lti	78.4	7.4	49.1	64.1	84.3	-4.6	86.4	-3.9	94.0	-4.3
lto	89.2	0.2	95.9	2.1	95.4	1.2	92.2	0.7	95.4	-2.0
ltz	89.2	4.0	92.9	-5.4	96.0	-1.5	89.0	2.8	98.0	0.3
lua	83.9	-3.5	93.0	3.7	91.4	1.6	84.7	-7.2	93.1	4.0
lue	85.4	6.5	94.0	8.3	94.2	0.3	90.8	-3.4	94.8	0.5
lug	75.3	-5.8	89.8	6.2	83.4	-5.8	72.5	4.3	85.6	3.9
lun	92.0	-9.0	95.1	-7.9	97.5	-0.6	96.9	2.2	98.6	0.3
luo	94.8	-5.0	98.2	-1.6	97.0	-2.0	95.4	-2.7	97.2	-2.7
lus	78.9	5.5	88.0	11.7	83.3	5.3	78.7	15.2	86.4	9.2
lvs	89.5	-10.8	95.8	-5.3	96.2	-1.1	94.2	-4.0	97.9	-1.2
lwg	61.4	17.7	83.5	9.0	79.7	1.3	67.7	5.4	79.4	6.7
lwo	67.0	2.4	76.7	13.4	83.4	-3.5	83.0	1.2	86.8	-3.2
lww	97.6	1.1	99.2	-1.2	97.6	2.1	96.6	5.7	98.3	1.3
lzh	76.1	3.9	71.8	-27.1	44.6	63.2	54.1	61.1	56.7	59.6
lzz	97.4	0.8	96.3	-4.0	97.0	-0.8	97.8	1.1	98.1	-0.5
maa	99.0	1.4	99.8	0.1	99.3	0.4	99.4	0.8	99.7	0.4
mad	80.8	0.5	91.0	12.5	89.2	9.1	86.4	12.2	88.1	0.9
maf	96.9	-2.5	99.0	0.1	98.2	-0.4	97.9	-0.7	97.9	-2.4
mag	67.0	-5.2	95.6	-7.4	98.2	-0.9	92.6	-7.4	99.1	-1.1
mai	50.0	-16.2	58.7	-50.1	73.8	-28.1	81.3	-13.0	91.1	-1.0
maj	98.1	0.1	99.7	0.3	98.7	1.2	98.1	2.3	99.1	0.1
mak	87.5	7.0	90.8	8.8	91.9	5.6	85.8	-3.4	91.7	3.1
mal	88.6	10.9	95.1	-1.7	98.1	0.6	97.6	1.4	99.1	1.2
mam	97.1	-2.3	99.5	-0.8	98.9	-1.0	97.7	-1.3	99.3	-0.4
maq	99.1	1.0	99.4	-0.3	99.1	0.6	99.1	0.6	99.5	0.0
mar	47.8	-32.2	78.2	0.7	92.0	-7.2	87.1	-7.0	95.3	-3.0
mas	77.4	-1.1	90.4	8.2	88.0	10.9	83.4	12.2	87.8	11.2
mau	99.9	0.1	99.6	-0.9	100.0	0.1	99.9	0.1	99.9	0.1
mav	98.7	-2.6	99.7	-0.5	99.6	-0.3	99.7	0.5	99.8	0.2
maw	90.0	5.9	97.5	1.8	94.3	2.2	91.2	2.5	94.2	0.3
max	56.9	-7.4	78.8	8.3	71.3	-10.8	66.9	-7.0	74.9	-10.0
maz	98.6	0.7	99.7	-0.4	99.0	0.1	99.0	0.5	99.3	-0.5
mbb	94.8	-5.3	98.7	0.0	97.4	1.4	96.0	7.1	98.1	3.3
mbc	98.5	1.0	98.8	-0.8	98.7	1.5	98.5	2.7	99.0	1.9
mbd	92.7	4.9	98.9	-1.9	98.5	-0.1	96.6	1.1	99.1	0.3
mbh	94.5	1.8	98.3	1.4	95.4	1.0	94.2	-0.2	96.0	1.0
mbi	95.5	1.0	99.4	-0.7	98.5	-1.8	96.4	-3.8	98.9	-1.8
mbj	98.6	0.1	99.9	-0.2	98.3	1.5	99.0	1.9	99.0	0.9
mbk	93.2	2.2	95.5	0.3	94.3	-1.9	90.1	4.0	93.1	-1.7
mbl	86.6	-21.5	86.6	-22.3	99.2	-1.1	99.7	-0.2	99.5	-0.5
mbs	94.8	-5.1	99.4	-1.1	98.7	0.2	97.8	0.4	99.4	0.4
mbt	96.3	-4.9	99.2	-1.2	98.3	-0.7	97.2	-2.8	99.5	-0.3
mbu	87.0	15.0	85.8	4.2	87.7	-1.9	87.5	-11.3	89.5	-10.2
mca	99.8	-0.3	99.5	-1.1	99.9	-0.2	99.8	0.3	99.9	0.2
mcb	97.9	0.0	97.2	-4.9	98.5	-1.4	97.3	-1.4	98.6	-0.5
mcd	96.0	-2.6	97.5	0.0	97.1	0.0	97.4	3.1	98.2	2.6
mcf	97.7	-4.5	97.9	-4.2	99.3	-1.0	99.1	-0.5	99.5	0.0
mch	99.6	0.1	99.2	-1.5	99.5	-0.5	99.5	0.2	99.6	0.2
mck	81.5	9.0	92.3	9.1	87.0	-2.1	79.9	-4.5	89.0	-0.3
mcn	92.8	2.1	96.7	4.2	94.7	3.6	93.0	1.6	95.8	3.8
mco	99.1	-0.1	99.4	-0.7	99.4	0.0	99.5	0.0	99.5	-0.3
mcp	99.0	-0.7	99.8	-0.5	98.5	1.8	98.5	1.7	98.9	0.4
mcq	97.0	-3.0	99.4	-0.9	99.1	1.2	99.1	1.4	98.8	1.4
mcr	100.0	0.0	99.7	-0.6	100.0	-0.1	99.9	0.2	99.9	0.1
mda	98.6	2.5	98.6	2.4	96.5	-0.7	96.9	-2.1	95.7	-5.0
mdf	65.6	-7.8	61.9	44.5	85.7	-7.3	77.7	5.9	85.7	-16.6
mdy	88.9	-17.4	67.5	-49.0	99.2	0.7	95.9	6.8	99.4	1.2
med	99.5	-0.9	99.9	-0.2	99.6	-0.7	100.0	0.1	100.0	0.0
mee	96.2	-2.7	99.2	0.0	97.4	0.1	97.3	0.3	98.3	0.1
meh	99.0	1.2	99.4	-0.2	97.8	-2.4	99.2	-0.4	99.5	0.0
mej	99.1	-1.3	98.8	-2.0	99.8	0.2	99.2	0.9	99.9	0.2
mek	97.5	0.3	97.6	-1.4	97.8	1.8	97.5	3.7	97.9	1.2
men	87.6	10.2	91.3	13.0	86.4	-6.0	88.5	3.8	89.1	-1.9
meq	96.3	1.7	99.1	-0.4	97.9	1.1	96.1	0.3	98.4	1.2
mer	75.2	1.4	88.1	6.7	84.3	10.9	76.5	-8.6	81.8	14.9
met	95.6	-4.0	99.0	-0.7	96.7	-2.5	94.5	-2.7	98.1	-0.8
meu	75.6	-8.6	91.5	-1.3	85.7	20.0	84.0	20.1	89.0	17.7
mev	97.1	0.6	99.2	0.7	97.9	0.4	97.0	1.4	97.6	0.3
mfe	83.3	-6.1	94.1	6.3	88.5	3.5	79.6	-7.1	87.5	-3.1
mfg	83.2	6.6	94.8	6.7	91.8	7.9	86.5	-0.1	91.2	4.7
mfk	91.6	0.6	95.4	5.2	92.7	-7.6	87.1	-14.6	89.5	-13.9
mfq	95.4	6.3	95.4	3.7	96.5	0.0	97.0	0.3	97.1	-2.0
mfv	81.6	2.4	83.2	-7.3	86.8	-1.2	80.3	-17.6	86.0	-10.3
mfy	97.3	1.6	98.7	1.7	97.6	0.5	98.0	2.3	98.5	1.5
mfz	99.1	-1.6	99.6	-0.8	99.7	-0.5	99.8	0.1	99.9	0.1
mgc	99.4	0.8	99.8	-0.2	99.0	0.3	99.2	0.2	99.7	0.2
mgd	99.1	-1.2	97.7	-4.6	99.7	0.1	99.6	-0.2	99.8	-0.1
mgg	96.1	4.8	97.5	2.7	97.4	2.6	96.7	1.8	97.6	2.7
mgh	78.2	6.1	94.3	4.7	94.3	1.6	89.6	0.5	95.9	3.0
mgo	98.0	2.2	98.4	2.9	97.1	2.3	97.7	3.2	97.9	2.5
mgp	92.6	4.5	96.4	2.6	96.1	4.3	94.4	1.6	95.3	1.0
mgr	72.7	-8.7	88.5	1.3	79.5	-5.3	74.5	4.4	84.3	6.3
mhi	98.8	-1.1	99.4	-1.2	99.5	-0.6	99.0	-0.2	99.7	-0.2
mhl	96.9	1.0	97.3	0.2	96.2	-1.0	95.2	-0.9	96.4	-1.7
mhr	86.6	-9.2	94.6	-5.8	96.0	-3.9	93.3	-3.9	96.5	-4.3
mhx	91.1	13.1	98.4	1.2	98.0	0.9	97.6	0.9	98.0	-0.6
mhy	89.1	6.9	93.4	1.8	92.1	2.5	87.9	-1.7	93.8	-0.7
mib	98.7	-0.4	99.8	0.2	99.2	-0.2	99.0	0.2	99.5	-0.5
mic	98.5	1.7	96.3	-4.5	98.6	-0.8	98.7	-1.0	98.8	-1.4
mie	97.8	-2.9	99.6	-0.8	99.4	-0.3	99.3	-0.1	99.7	0.3
mig	99.1	-1.5	99.8	-0.4	100.0	-0.0	99.9	0.1	100.0	-0.1
mih	98.3	-0.9	99.8	0.0	99.4	0.4	98.7	-0.3	99.5	0.8
mil	99.5	-0.2	99.7	-0.7	99.3	-0.4	99.5	0.3	99.8	0.1
mim	97.1	-0.2	99.2	-1.2	97.0	-4.4	96.0	-0.5	97.1	-4.0
min	79.0	-8.1	93.0	-2.4	94.4	-4.9	91.8	-0.5	98.2	0.8
mio	88.3	-20.2	99.7	-0.7	99.4	-0.3	98.4	-0.9	99.2	-0.4
mip	96.3	-4.4	98.9	-1.4	98.3	1.4	98.2	0.8	99.4	0.4
miq	89.7	-1.0	96.1	5.5	94.2	2.2	89.7	-4.1	95.6	2.9
mir	99.5	-0.5	99.4	-0.8	99.6	-0.3	99.6	0.2	99.8	0.2
mit	98.9	-0.5	97.6	-3.8	99.4	0.5	99.3	0.7	99.5	0.4
mix	94.9	-1.0	99.1	0.1	97.4	1.4	96.2	-1.6	97.7	-0.6
miy	99.4	-1.3	99.8	-0.5	99.8	-0.3	99.9	0.0	100.0	0.0
miz	98.9	-1.1	99.6	-0.6	98.6	0.8	99.1	0.6	99.4	0.1
mjc	97.7	-1.5	99.1	-1.3	98.9	0.0	98.5	0.8	99.4	0.4
mjl	73.1	-7.0	96.7	-5.2	97.0	-3.4	95.4	-3.9	97.6	-2.7
mjv	92.5	-5.6	95.8	2.9	98.8	0.7	98.0	1.2	99.2	-0.2
mjw	88.5	0.4	93.9	5.7	93.7	3.4	87.3	-3.8	92.4	-2.9
mkd	67.6	-16.0	71.0	-16.5	87.8	0.3	82.7	-5.0	93.5	1.8
mkj	97.7	-1.2	98.0	-2.7	97.5	-2.4	96.7	-3.9	97.5	-2.6
mkl	98.8	0.3	99.0	-1.1	98.8	1.5	99.0	0.9	99.6	0.1
mkn	74.6	16.6	94.3	5.8	88.6	16.0	78.0	30.2	91.4	14.0
mks	99.5	-0.3	100.0	-0.1	99.7	0.1	99.7	0.4	99.9	-0.0
mle	98.6	-1.9	98.4	-3.1	99.0	-0.9	99.2	-0.5	99.3	-0.7
mlh	99.1	-1.1	99.0	-1.7	99.5	-0.3	99.4	0.0	99.4	-0.7
mlk	97.6	1.6	99.0	-1.6	97.9	-0.7	98.7	-0.8	99.5	0.1
mlp	96.7	-5.0	99.1	-1.1	98.8	-1.0	99.2	0.8	99.5	0.0
mlt	81.3	-0.2	83.2	7.5	83.0	-1.1	85.5	-4.7	92.4	-3.3
mlu	96.3	-0.2	99.2	-1.1	97.7	-0.7	96.0	-4.0	98.4	-1.7
mmn	97.3	-2.6	99.4	-1.2	99.3	-1.1	98.9	0.9	99.8	-0.1
mmo	83.3	-27.1	47.6	-68.4	98.5	-1.2	98.9	0.2	99.3	0.3
mmp	92.1	-6.4	95.6	0.2	96.2	2.5	96.3	3.6	97.2	3.9
mmx	94.6	-4.8	99.5	-0.2	98.7	-0.7	98.5	1.0	99.3	1.1
mmy	84.1	10.9	93.7	9.1	91.8	-5.1	87.2	-8.7	89.4	-12.8
mna	95.6	-6.9	99.1	-1.6	98.8	-0.3	98.0	0.9	99.0	-0.5
mnb	97.3	-4.8	98.7	-2.3	99.8	-0.3	98.9	1.4	99.7	-0.5
mnf	98.4	2.3	98.7	1.7	97.4	0.8	97.1	1.2	97.5	2.1
mni	50.0	65.4	49.3	66.5	96.9	0.7	89.6	16.5	97.9	-0.3
mnk	93.6	5.1	96.3	6.0	93.4	1.6	89.1	-3.2	92.9	-1.5
mnw	94.4	6.2	97.6	0.3	98.3	0.4	98.2	1.0	98.6	-0.5
mnx	92.3	4.1	95.3	7.0	94.1	4.5	93.8	6.3	95.0	5.7
moc	92.3	10.1	99.0	-0.2	94.7	6.4	91.0	10.2	96.6	2.8
mog	82.7	22.3	83.3	26.7	83.7	19.4	79.1	7.1	84.6	16.6
moh	97.1	4.3	95.3	-0.9	96.1	1.6	97.0	3.0	97.2	1.4
mop	97.3	-4.1	99.4	-0.6	99.1	-0.3	97.9	2.6	99.0	0.3
mor	99.8	-0.4	99.9	-0.3	99.7	-0.4	100.0	0.1	100.0	0.1
mos	92.6	7.5	94.8	7.1	93.5	6.8	95.0	1.0	96.2	1.0
mox	97.0	3.5	98.7	1.8	96.8	1.8	96.7	5.3	98.1	2.3
moy	95.5	3.5	96.4	-0.9	93.9	-3.2	94.4	0.3	96.1	-0.6
mpd	96.1	-5.0	98.4	-2.9	99.2	-1.2	99.3	0.1	99.5	-0.1
mpe	96.6	1.9	97.3	0.1	95.6	-2.0	94.6	-1.3	97.1	-1.2
mpg	93.8	4.1	95.8	7.1	92.7	2.2	90.6	0.2	92.5	-3.2
mpj	96.0	-2.4	97.8	-4.3	98.9	-1.3	96.4	-3.4	99.6	-0.6
mpm	95.9	-2.5	99.6	-0.9	98.1	-1.6	96.0	-4.8	98.6	-1.5
mpp	97.8	0.2	99.1	-1.1	98.0	0.4	98.5	0.3	99.3	-0.4
mps	87.3	-16.0	98.6	0.8	97.8	1.7	96.4	5.6	98.3	2.9
mpt	98.4	-2.6	99.3	-1.3	98.9	-0.4	99.3	0.3	99.5	0.2
mpx	94.5	-6.3	98.8	-1.8	98.5	0.3	95.7	0.0	98.3	0.0
mqb	98.2	1.3	99.6	0.2	98.3	2.5	97.5	3.2	98.8	2.1
mqf	98.2	-3.5	98.4	-3.1	100.0	-0.1	99.9	0.1	99.9	-0.2
mqj	79.0	-17.4	93.5	-7.8	94.0	-3.7	89.2	0.8	91.4	-10.4
mqn	97.5	-2.0	99.0	-1.6	99.0	0.2	98.2	-1.1	99.4	-0.3
mqy	91.6	9.0	96.0	6.0	91.8	2.9	93.1	5.1	93.0	-1.0
mrg	97.0	3.1	97.2	-0.6	98.1	-0.7	97.0	-2.6	98.8	0.0
mrh	93.1	4.9	95.1	3.0	95.3	3.8	91.8	-5.1	96.5	0.3
mri	89.7	-3.0	93.4	-2.3	92.3	-4.1	90.7	-5.2	96.3	-1.0
mrj	91.6	12.3	89.2	-2.5	94.1	7.2	94.5	7.7	95.8	6.3
mrn	95.3	4.7	97.7	3.0	96.9	-0.3	96.1	3.9	98.3	1.4
mrw	95.7	-7.1	98.9	-2.2	99.3	-0.3	98.5	-0.7	99.7	-0.1
msb	86.1	-5.8	94.4	-8.3	95.7	-3.9	92.7	-0.8	97.5	-2.8
msc	88.7	-9.9	97.8	-3.2	96.7	1.2	92.2	-5.2	96.0	5.0
mse	89.2	9.8	93.5	9.6	89.2	6.7	85.1	0.7	93.3	2.5
msk	90.3	-5.5	98.7	-2.2	96.5	-3.3	94.0	-5.2	98.5	0.4
msm	89.7	-7.1	98.2	1.9	95.9	-0.4	91.2	-4.0	97.2	-0.6

Table 7: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
msy	96.3	1.9	98.1	2.1	97.3	1.5	97.5	4.1	98.1	2.3
mta	98.9	0.0	99.6	-0.3	99.4	-0.4	99.7	0.0	99.7	0.1
mtg	96.2	1.1	96.8	0.6	97.4	2.5	96.5	4.6	97.5	3.6
mti	95.5	-2.0	98.6	-0.4	97.1	2.4	93.8	-0.1	97.6	0.4
mtj	99.5	-0.5	98.4	-2.6	99.4	-0.6	99.1	-1.0	99.8	0.2
mto	99.4	-0.3	99.7	-0.2	99.4	0.6	99.5	0.7	99.8	0.0
mtp	98.6	-0.5	99.3	-0.6	98.7	-0.7	98.5	0.6	99.2	0.4
mtq	90.3	-0.2	90.3	-5.4	95.6	-1.8	93.0	1.6	96.6	-0.3
mtr	71.8	22.9	83.9	20.3	89.7	5.6	85.1	10.5	88.9	2.0
mtt	97.1	-2.7	99.4	-0.1	97.8	-2.9	97.8	-2.1	98.9	-1.1
mua	87.7	20.1	90.8	14.1	89.2	15.7	90.9	12.0	93.2	10.0
mug	95.9	3.8	97.9	3.0	96.7	4.2	95.7	5.0	95.9	4.6
mup	76.7	-6.1	96.6	-6.3	98.9	0.1	98.0	1.9	99.5	0.2
mur	95.2	-1.2	98.2	0.4	97.5	1.4	97.1	1.4	98.4	1.1
mux	92.5	1.7	97.7	1.0	96.0	1.1	94.2	2.9	96.9	-0.7
muy	98.4	-0.5	99.4	-0.8	98.3	0.6	97.2	2.2	98.9	1.5
mva	97.2	-4.2	98.5	-1.9	98.4	-1.2	98.0	-0.8	98.7	-0.9
mvn	93.4	5.5	98.6	0.5	95.6	0.6	93.5	0.9	96.8	1.7
mvp	58.5	-21.2	76.7	-17.5	78.0	-17.4	80.0	-11.1	79.0	-11.5
mwc	91.2	-13.7	97.7	-3.2	95.6	-5.4	94.8	-4.4	97.3	-4.1
mwe	90.8	3.9	96.8	-2.5	95.8	0.7	93.4	5.6	96.5	1.7
mwf	99.8	-0.5	97.5	-4.8	99.8	-0.2	99.9	0.0	100.0	-0.1
mwh	95.6	3.6	98.2	2.5	95.5	-0.2	93.6	-1.3	93.9	-6.4
mwl	83.8	-7.2	94.2	-2.9	93.1	-3.6	87.3	-8.2	96.6	-1.3
mwm	91.1	14.4	91.5	14.7	93.1	4.0	92.4	5.5	95.1	6.6
mwn	78.8	-14.0	91.5	-3.3	85.6	-8.1	77.2	-11.3	86.1	-5.4
mwp	98.8	-1.5	99.7	-0.4	98.9	-1.1	98.3	-0.3	98.6	-2.2
mwq	83.2	18.5	87.0	18.6	82.6	3.0	79.2	-13.0	83.2	-2.3
mwv	76.0	17.5	80.4	16.2	70.0	-17.9	66.1	-16.5	68.7	-19.4
mww	94.0	6.8	95.5	4.3	95.0	0.1	88.4	-9.2	95.6	-0.7
mxb	96.9	0.5	98.9	-2.0	99.4	0.1	97.6	-3.0	99.6	-0.2
mxp	99.1	1.5	99.5	-0.2	99.2	0.7	99.2	0.5	99.7	0.5
mxq	99.3	0.8	99.6	0.4	99.5	0.6	99.4	0.5	99.7	0.6
mxt	96.9	-1.6	99.6	0.1	98.4	0.1	97.9	-0.8	98.6	-1.3
mxv	74.0	36.8	94.7	9.5	96.2	4.7	95.3	6.8	98.4	1.8
mxx	87.7	16.1	91.2	15.0	91.5	3.5	89.3	-1.7	90.3	-0.1
mya	94.3	-0.0	95.8	3.2	96.0	1.2	96.3	1.9	96.5	-0.6
myb	94.9	-0.8	97.3	4.1	94.1	7.4	95.6	4.8	96.8	4.5
mye	97.5	3.6	98.6	1.7	96.6	5.1	97.5	4.2	98.2	3.3
myk	95.5	3.3	98.8	1.0	97.0	0.4	95.0	1.3	96.7	-0.4
myu	89.8	-17.0	95.2	-7.6	98.9	0.1	99.0	0.8	99.0	0.2
myv	72.0	-10.5	75.9	-20.2	89.3	-4.0	80.6	-11.4	94.1	0.5
myw	97.1	-3.8	98.6	-2.3	98.0	-2.2	97.7	-1.8	98.3	-1.9
myx	84.6	1.5	92.1	5.2	91.5	5.8	86.2	2.9	90.3	4.9
myy	93.3	-6.1	98.2	-3.6	98.7	-0.6	96.4	-0.4	98.7	-1.8
mza	95.6	-0.4	98.3	0.4	97.2	0.6	97.0	1.3	98.4	2.5
mzh	97.2	3.4	97.9	1.6	96.0	0.0	96.3	-1.4	97.9	-0.7
mzi	98.0	-1.6	99.7	-0.4	99.3	-0.7	98.2	-2.0	99.2	-0.8
mzj	80.6	-5.0	83.6	-19.4	87.5	-7.0	84.5	-4.0	87.9	-11.9
mzk	99.3	-1.3	100.0	-0.1	99.9	-0.1	99.9	0.2	99.9	-0.2
mzl	99.6	-0.2	99.8	0.1	98.4	-2.0	99.4	0.0	99.6	0.0
mzm	96.5	2.9	98.6	2.1	96.9	1.5	93.3	-5.1	97.0	0.6
mzn	23.7	11.8	22.8	15.3	68.9	32.4	5.3	27.4	25.8	62.6
mzr	88.6	19.3	88.2	17.5	88.5	18.6	88.7	19.5	88.9	19.7
mzw	90.8	8.6	94.3	9.7	90.0	0.9	88.8	0.8	89.9	1.3
nab	99.9	-0.2	98.6	-2.8	100.0	-0.1	100.0	0.0	100.0	0.0
naf	98.2	-1.3	99.1	0.1	99.0	0.1	98.8	0.3	99.3	-0.0
nag	93.3	-8.3	98.5	-1.4	97.8	2.7	96.9	3.7	98.7	1.8
nak	97.1	-0.2	99.2	-0.3	98.2	-0.9	97.0	-1.3	97.9	-2.0
nan	73.8	36.2	73.7	37.5	78.0	7.4	83.4	9.0	89.6	-4.4
nap	56.2	39.3	69.2	32.0	66.1	36.1	42.4	57.1	88.5	15.9
naq	96.3	5.7	95.3	3.1	91.8	-2.3	93.6	-2.1	95.9	-0.6
nas	99.4	-0.6	99.0	-2.0	99.0	0.0	99.6	-0.3	99.6	-0.3
nav	96.2	0.6	96.2	0.8	95.3	-2.5	96.9	0.2	98.2	1.0
nbc	95.8	4.5	97.4	2.3	92.6	-3.5	89.6	-4.5	92.7	-6.9
nbe	93.1	8.7	92.8	4.7	93.9	3.4	91.4	-3.2	95.6	-1.4
nbl	64.2	6.9	90.1	-5.9	89.5	-1.8	79.6	-4.0	90.7	-2.5
nbq	92.6	9.1	98.3	0.2	96.5	1.3	94.5	3.9	96.7	3.4
nbu	92.6	9.7	91.7	6.1	91.6	6.5	89.5	3.8	92.6	0.3
nca	98.2	1.1	98.4	-2.7	97.7	0.2	98.0	2.8	98.7	0.7
nce	98.7	-1.8	99.4	-1.1	99.0	-1.0	99.6	0.6	99.8	0.0
nch	78.6	-14.1	88.9	-3.5	86.4	-4.5	82.0	-10.7	87.3	0.3
ncj	92.9	1.2	98.3	-2.1	97.1	-2.0	94.8	-2.1	98.0	-0.8
ncl	91.7	-6.4	97.0	-3.5	96.7	-2.6	95.0	3.5	98.4	-0.2
nct	90.3	4.0	95.8	5.8	94.2	-1.4	91.6	-1.6	93.7	-4.7
ncu	99.3	0.8	99.3	0.0	99.3	-0.2	99.0	0.0	99.4	0.2
ndc	77.9	0.8	93.0	3.2	91.2	-1.7	84.9	-6.2	93.8	-3.8
nde	61.1	-9.8	83.4	-14.5	82.3	-8.8	72.8	-19.3	87.8	-9.0
ndg	74.3	-8.5	96.2	-3.2	92.3	-0.2	84.6	7.4	94.0	-0.3
ndi	94.2	9.0	96.3	6.8	93.7	-5.1	94.4	-4.3	96.2	-1.9
ndj	82.2	-6.8	95.8	-7.5	96.2	-4.0	90.8	-1.5	94.7	-5.0
ndo	84.6	-10.5	95.1	0.5	93.6	-0.1	89.8	-0.1	95.3	2.2
ndp	95.2	6.5	97.4	2.4	95.4	-0.3	93.6	-0.6	94.3	-3.6
nds	43.0	-11.2	59.9	19.4	66.7	-0.3	39.2	-28.1	68.2	-7.5
ndx	95.6	6.6	97.8	3.0	93.0	-3.1	86.3	5.7	90.3	-4.7
ndy	99.1	-1.2	98.5	-3.0	99.7	0.0	99.5	0.1	99.7	-0.5
ndz	98.2	-2.0	97.0	-5.6	98.6	0.5	98.3	1.8	98.3	-0.0
new	66.7	20.9	93.5	6.7	93.8	-4.9	95.0	1.5	96.6	-1.6
ney	93.3	9.4	98.8	-0.4	98.5	1.5	98.4	1.3	98.2	0.0
nfa	94.2	8.3	97.1	2.4	93.0	10.4	93.4	9.9	96.8	4.9
nfr	96.8	3.7	96.7	-1.2	98.4	1.3	98.1	1.2	99.3	0.6
nga	95.9	-1.0	97.8	2.2	96.9	2.4	96.4	0.2	98.2	-0.4
ngb	95.5	4.1	97.6	3.5	93.2	-0.5	93.8	0.7	96.3	0.0
ngc	90.6	-0.3	97.0	3.0	92.8	-5.5	91.4	-2.0	92.8	-7.3
nge	96.7	3.4	97.7	3.9	96.8	1.7	97.4	3.0	97.8	2.2
ngj	92.9	11.0	94.9	9.0	91.0	2.9	91.9	6.5	92.0	1.1
ngl	87.0	-4.3	92.4	5.4	93.3	-7.2	91.2	-9.9	94.6	-7.2
ngn	95.5	2.6	97.3	3.2	95.8	2.5	94.8	4.6	96.0	1.8
ngp	77.6	14.2	94.7	-1.2	91.4	-0.4	82.5	2.3	91.0	1.4
ngu	91.3	-0.2	94.7	-6.1	96.4	-1.4	93.4	5.6	97.1	-1.2
nhe	65.3	-1.0	77.0	-6.0	74.3	-1.5	67.1	-0.9	77.2	2.7
nhg	97.7	3.0	98.9	-0.6	98.4	1.5	96.8	2.7	98.9	1.0
nhi	97.6	-0.5	98.9	-0.8	97.8	-0.6	95.7	1.2	98.6	1.6
nho	99.9	0.1	99.4	-0.9	99.6	0.1	99.9	0.3	98.7	-1.0
nhr	96.8	4.9	98.7	2.4	97.1	1.4	97.6	1.7	98.3	1.4
nhu	98.2	-1.4	99.6	-0.7	98.8	0.4	98.4	2.0	98.9	-0.4
nhw	61.5	16.7	72.5	16.7	71.6	14.9	64.5	6.1	72.0	-5.5
nhx	97.7	2.2	98.6	1.2	98.2	2.0	97.6	1.8	98.2	0.3
nhy	96.0	1.7	99.1	-0.6	98.6	-0.3	96.3	-3.2	98.6	-0.6
nia	84.4	6.3	85.9	20.4	93.8	-2.4	93.7	2.9	94.8	-0.3
nif	99.5	-0.6	99.7	-0.5	99.5	-0.6	99.6	0.1	99.8	0.2
nih	90.2	11.2	94.7	2.6	90.9	-8.8	88.7	-12.9	93.1	-5.1
nii	99.1	-1.6	99.5	-0.9	99.5	-0.7	99.9	0.1	99.9	-0.0
nij	74.9	22.6	83.5	26.5	77.9	13.5	59.8	-17.9	75.7	5.9
nim	95.1	3.9	96.6	1.5	92.6	-3.8	94.8	1.4	94.9	-1.3
nin	97.5	1.6	99.3	0.5	97.6	1.6	95.8	0.7	98.4	-0.1
niy	99.2	-0.6	99.1	-1.9	99.9	0.1	99.8	0.2	99.9	0.1
njb	93.4	4.7	97.1	1.4	94.0	-0.2	92.9	0.4	93.8	-2.7
njh	91.0	9.1	92.3	-5.3	89.5	6.7	88.5	11.6	91.7	4.0
njm	89.5	1.0	91.2	-1.4	89.1	-6.0	86.4	-8.8	90.9	-3.7
njn	88.6	9.7	92.2	9.3	91.1	6.1	85.0	-4.9	92.5	1.8
njo	90.5	5.8	93.2	9.6	92.9	1.6	90.0	4.9	94.3	-2.4
njz	96.8	1.7	97.9	-0.1	97.6	1.8	97.2	-1.0	98.5	1.2
nka	77.8	13.9	91.3	9.2	89.6	3.6	84.3	-0.8	89.2	5.6
nkf	95.9	-6.9	97.4	-5.0	99.1	-0.4	98.5	0.7	99.7	-0.1
nki	61.2	11.7	78.3	6.4	79.5	-6.7	70.1	-10.7	76.3	-12.6
nko	98.0	-2.8	99.2	-1.4	99.2	-0.6	98.8	-1.0	94.1	-10.0
nkr	79.2	16.7	82.7	26.9	79.7	9.6	72.5	-9.4	79.7	0.0
nla	97.8	3.3	98.8	1.8	98.5	1.0	97.4	0.5	98.8	1.4
nlc	82.1	13.8	79.5	3.2	83.4	15.9	71.5	-12.8	84.9	19.5
nld	73.1	-5.2	77.3	-15.6	74.4	-20.1	40.8	-57.1	60.2	-47.4
nlg	89.5	-5.6	97.1	1.9	93.8	2.4	90.2	5.0	95.8	5.4
nlk	56.8	17.8	68.0	48.3	58.9	9.3	51.8	2.6	56.6	-0.1
nlx	77.8	-15.0	96.0	-6.9	98.8	-0.8	95.3	-1.8	98.5	0.1
nma	94.8	1.7	97.4	1.2	95.7	1.4	94.9	-0.5	96.0	-0.4
nmf	80.9	19.9	94.0	8.4	91.4	3.0	89.9	0.5	94.5	3.0
nmh	94.7	2.9	96.2	-0.9	93.4	-4.5	90.5	-7.9	92.7	-6.2
nmo	89.2	7.6	92.3	8.6	91.0	2.2	89.0	1.7	91.9	2.4
nmz	99.4	-0.2	99.0	-2.0	99.5	-0.8	99.7	0.5	99.9	-0.0
nna	96.3	-1.4	98.4	-0.4	95.7	-3.1	92.9	-2.3	98.0	-0.6
nnb	97.5	2.5	98.6	0.8	97.7	0.3	98.3	1.0	99.0	-0.4
nnd	42.5	-26.1	39.9	-62.0	67.8	-25.9	67.8	-27.4	78.3	-17.9
nng	92.4	2.4	96.5	4.0	94.4	0.6	93.9	1.4	95.9	1.4
nnh	97.8	0.6	98.0	-0.2	96.9	-1.8	97.4	-0.6	98.3	-0.4
nnl	88.0	7.8	93.0	9.3	91.2	4.6	88.6	3.2	92.1	0.4
nno	63.4	-6.8	86.4	1.6	82.5	-2.8	69.5	12.8	90.9	1.3
nnp	91.7	10.9	93.5	8.4	92.2	1.8	87.4	-4.9	93.6	-0.6
nnq	88.1	-9.4	97.6	-3.4	95.0	-3.5	90.3	-5.4	95.2	-3.3
noa	97.6	2.0	98.5	2.2	97.5	1.3	98.1	3.0	98.3	2.5
nob	63.6	-3.6	80.1	16.8	80.4	-6.3	66.0	-20.2	90.7	-4.1
nog	74.8	-13.0	83.0	-7.0	91.4	-6.8	84.1	-11.2	92.4	-5.0
noi	84.5	2.0	98.4	-2.0	99.2	0.7	96.9	-0.5	99.4	0.4
nop	98.7	-1.7	99.4	-1.0	99.6	-0.1	99.2	0.3	99.5	-0.3
not	97.5	-0.9	98.3	-1.7	98.5	0.7	97.3	-0.9	98.7	-0.1
nou	95.4	1.0	98.4	1.8	97.5	1.2	96.7	1.5	98.5	2.4
nov	46.4	0.0	70.8	29.1	60.5	2.9	50.2	13.7	69.5	-2.7
nph	92.3	5.5	92.6	1.7	91.1	0.8	92.3	9.0	94.6	5.5
npi	49.5	-9.5	74.5	-22.8	85.6	-9.8	74.1	-31.2	92.4	-4.7
npl	93.9	0.3	97.7	1.4	96.7	-2.4	94.3	-1.9	97.8	-0.6
npo	93.0	6.9	95.0	7.3	93.6	2.3	92.3	3.1	93.4	0.3
npy	82.5	3.7	93.1	7.9	87.1	3.3	81.7	-1.1	84.5	-14.5
nqo	100.0	0.0	99.5	-0.5	99.0	0.4	99.9	0.0	100.0	0.0
nre	94.4	6.4	95.3	5.7	94.6	5.5	93.9	3.3	94.7	5.2
nri	91.4	8.9	93.0	6.0	89.2	14.9	90.4	16.0	92.8	10.3
nsa	91.8	8.2	93.5	8.0	91.0	4.8	89.2	4.2	92.8	1.9
nse	77.2	14.7	91.2	6.0	83.8	-6.9	80.2	1.4	88.2	-8.3
nsk	90.9	3.5	78.7	-15.5	91.2	13.6	92.6	12.6	95.4	7.4
nsm	92.9	3.1	95.7	4.8	94.0	2.7	93.8	2.7	96.4	0.3
nsn	95.0	-5.3	98.4	-1.8	96.8	-3.5	94.5	-6.1	96.9	-3.4
nso	85.0	-0.4	89.9	1.7	88.7	-4.1	85.8	-5.4	93.3	-2.1
nss	96.1	-6.5	99.3	-1.0	98.7	-1.5	98.7	-0.6	99.2	-0.3
nst	94.2	9.5	95.4	5.1	93.2	6.4	93.8	5.8	96.3	2.6
nsu	96.2	-1.5	98.2	-2.2	98.2	-0.5	96.5	-3.0	99.5	-0.1
ntj	97.2	-1.9	92.8	-13.1	98.2	0.7	96.8	-3.6	99.0	-0.3
ntk	85.5	16.6	93.9	8.4	92.6	-5.0	87.9	-7.7	92.3	-4.6
ntm	99.6	-0.9	99.4	-1.3	100.0	-0.1	99.9	-0.1	99.9	-0.0
ntp	99.7	-0.2	96.3	-7.1	99.5	0.5	99.6	0.8	99.7	0.6
ntu	99.1	-0.2	98.9	-1.6	99.1	-0.8	99.3	0.1	99.4	-0.2
num	82.9	12.8	87.8	18.9	90.6	-0.3	88.1	-9.1	93.0	-2.8
nus	95.9	6.3	97.4	2.9	96.9	2.7	97.7	0.4	98.4	0.3
nut	92.0	8.5	92.0	-0.7	93.7	-1.8	93.9	1.0	95.8	2.0
nuy	99.4	-0.4	98.9	-1.6	93.7	-10.4	91.7	-13.5	95.8	-6.7
nuz	90.7	3.3	93.8	5.4	95.5	-2.4	92.0	-5.9	97.1	-0.2
nvm	96.9	-3.9	99.2	-1.6	99.8	0.4	99.7	0.4	99.7	0.1
nwi	95.3	-7.5	99.5	0.1	99.0	0.5	97.6	0.1	98.7	0.7
nxd	97.4	-2.6	97.6	-4.0	97.6	-2.2	96.2	-1.8	97.3	-3.8
nya	78.2	-5.4	94.6	-2.3	94.1	0.2	86.9	7.3	96.4	0.7
nyb	85.6	1.8	92.9	-7.3	92.9	-1.2	89.8	0.5	94.7	2.5
nyd	68.0	-4.4	85.3	4.3	79.2	2.0	69.1	0.2	81.0	-0.6
nym	81.3	-16.6	98.0	-3.1	95.1	-5.4	86.5	-11.4	94.9	-5.7
nyn	77.5	-1.9	95.0	-3.5	92.3	-0.7	84.6	-6.6	93.5	1.5
nyo	68.3	16.8	90.3	8.7	88.9	1.1	79.8	6.1	90.1	9.0
nyu	81.8	-4.1	92.4	-11.8	92.9	-7.7	90.1	-0.4	90.2	-12.7
nyy	98.9	0.5	99.0	-0.7	98.2	-0.9	99.3	0.6	99.6	0.4
nza	98.2	1.7	99.0	1.1	97.6	2.6	97.7	1.4	98.0	0.6
nzi	91.1	5.2	94.7	7.7	94.1	0.8	88.7	-3.4	95.1	4.6
nzm	96.9	2.6	97.4	2.0	97.3	1.7	96.5	-1.3	98.3	-0.5
obo	97.5	-0.7	99.0	-0.7	97.8	0.7	96.9	2.2	98.4	2.4
oci	69.3	7.3	82.1	21.3	81.9	-6.0	62.2	-30.3	85.8	-11.2
odu	98.1	-0.5	99.2	-0.7	98.6	-0.2	98.2	-0.7	99.1	0.2
ojb	46.5	63.5	52.6	64.2	74.6	-4.9	80.9	3.8	81.5	-0.5
okr	92.8	-0.2	96.7	-0.3	97.3	1.4	95.9	-0.1	98.5	-0.8
oku	95.6	1.3	98.1	2.1	96.1	2.5	95.4	1.2	96.6	2.5
okv	98.4	-2.3	99.9	0.0	99.2	-0.5	98.8	-0.9	99.7	-0.4
old	84.2	5.4	92.8	4.5	90.5	2.9	86.4	11.1	89.0	12.3
olo	88.9	-16.2	90.8	-15.0	94.4	-8.1	92.7	2.1	96.0	-5.5
omb	94.8	2.1	98.6	-0.7	96.2	-1.8	92.8	-2.1	96.6	-1.4
omw	97.6	3.1	96.1	-4.3	98.0	2.3	97.0	5.2	98.1	3.1
onb	99.4	-1.3	99.1	-1.8	99.9	-0.2	100.0	-0.1	100.0	-0.0
ong	96.1	-6.9	98.9	-1.7	99.2	-0.8	99.3	0.3	99.5	-0.8
ons	99.0	-1.6	99.6	-0.9	99.5	-0.6	99.5	0.5	99.7	-0.1
ood	98.1	-1.0	99.1	-0.5	98.9	0.1	98.6	1.1	99.0	0.9
opm	96.0	-5.7	98.4	-2.8	99.2	0.2	98.5	1.8	98.4	1.7
orc	99.3	-1.3	99.4	-1.2	99.9	0.1	99.9	0.0	99.9	-0.1
oro	95.3	-3.5	96.7	-2.7	96.9	-1.3	94.2	-0.5	95.4	-3.7
ory	90.8	6.4	96.7	2.1	95.1	-4.6	94.0	-4.2	96.8	-4.2
oss	91.9	9.4	93.1	4.5	95.4	2.0	95.7	6.7	97.2	3.8
ota	80.7	-6.8	83.1	-13.5	92.1	-0.6	80.3	-18.3	93.2	-3.0
otd	66.5	-26.9	73.5	-22.4	82.7	-5.9	81.1	-5.0	87.2	1.1
ote	98.3	2.6	99.7	0.3	98.7	0.8	98.2	0.2	99.3	0.1
otm	99.6	-0.4	99.5	-1.0	99.5	-0.2	99.6	-0.1	99.8	0.1
otn	99.7	-0.3	99.9	-0.1	99.8	0.2	99.7	0.5	99.8	0.2
otq	98.7	-0.3	99.4	-0.1	98.9	-0.7	98.7	0.1	99.3	0.0
otr	96.6	4.5	97.2	4.7	98.5	-0.9	98.9	0.2	99.2	0.2
ots	98.5	-0.2	99.8	-0.2	99.1	0.5	99.3	0.3	99.9	-0.1
oym	99.6	-0.7	99.8	-0.3	99.8	-0.3	99.9	-0.1	99.9	-0.1
ozm	95.3	5.9	97.1	4.2	96.4	3.6	94.6	5.4	95.5	3.5
pab	98.2	-3.3	98.7	-2.1	99.8	-0.2	99.5	-0.6	99.4	-1.0
pac	98.0	3.0	98.6	1.5	97.9	0.7	98.1	0.9	98.3	-0.4
pad	97.5	-3.0	98.1	-3.4	98.3	0.1	97.5	-0.3	98.4	-0.1
pae	98.8	-0.5	98.2	-3.1	99.5	-0.3	99.5	0.3	99.7	0.1
pag	66.0	5.6	80.1	26.3	77.6	-2.4	72.7	-6.4	85.0	-4.7
pah	97.6	-1.7	99.5	-0.7	98.9	-0.0	98.0	-0.1	99.0	-0.7
pam	74.9	-5.1	89.8	11.9	88.0	-8.5	87.5	-2.4	95.8	1.0
pan	99.9	0.2	99.4	0.5	99.6	0.3	99.8	0.3	99.9	0.1
pao	97.1	-5.3	99.2	-1.5	99.6	-0.2	99.1	1.0	99.7	0.5
pap	60.1	11.1	73.1	35.9	65.9	8.4	54.1	-6.4	73.3	12.4
pau	95.2	-8.8	98.7	-2.3	99.6	-0.4	99.8	0.0	99.7	-0.1
pav	98.4	0.7	98.6	0.1	97.4	1.8	97.3	2.1	98.2	0.0
pbb	99.5	-0.4	97.9	-3.8	99.6	0.1	99.5	0.2	99.8	0.3
pbc	97.8	1.3	99.1	-0.2	96.8	-0.2	98.2	2.2	98.9	1.3
pbi	99.6	-0.5	99.5	-0.9	99.7	-0.7	99.8	0.0	99.9	-0.2
pbt	90.5	-13.3	90.6	-9.7	38.3	70.2	9.7	67.7	48.9	65.4
pbu	75.0	24.5	89.3	-1.1	94.5	6.6	72.5	-27.3	82.3	-25.4
pcd	49.6	-43.2	68.6	-31.2	83.6	1.7	53.4	44.7	69.7	41.8
pce	98.1	-3.7	99.9	-0.3	99.9	0.1	99.7	0.2	99.7	0.6
pck	64.0	4.9	84.8	13.1	81.8	8.0	72.9	11.4	79.8	9.8
pcm	74.1	-37.9	97.2	-0.5	99.1	-0.1	98.3	-2.3	99.7	-0.3
pdc	70.7	23.9	74.1	18.0	83.0	6.5	74.1	4.3	79.8	-15.4
pdt	89.0	7.5	92.9	6.5	92.1	-0.2	79.1	-19.6	92.9	-5.1
peg	90.8	-4.2	98.2	-2.7	99.2	-0.4	97.9	-1.4	99.2	-0.5
pes	40.3	4.7	18.3	71.9	59.3	3.2	43.9	-4.1	70.9	14.3
pez	64.2	-20.8	67.9	-48.4	61.5	-4.0	49.5	21.5	67.5	-24.0
pfe	89.9	9.6	93.2	12.8	92.5	10.4	91.6	9.1	92.8	8.7
pfl	79.8	-10.4	84.6	-19.3	89.1	-4.0	81.5	3.7	89.7	4.7
pgg	65.8	2.5	90.8	5.7	92.7	-4.2	82.3	-16.8	92.1	-3.3
phr	91.5	7.8	94.3	-8.3	97.3	-1.1	91.8	-5.0	98.0	-1.2
pib	94.3	-8.6	97.0	-5.6	97.1	-3.2	97.3	-1.4	98.1	-2.2
pid	100.0	0.0	98.9	-2.1	100.0	-0.0	99.9	0.1	100.0	0.0
pil	97.9	-4.0	99.4	-1.3	98.9	-1.6	99.4	-0.1	99.8	-0.2
pio	99.2	-1.5	99.0	-1.9	99.7	-0.7	99.9	-0.3	99.9	-0.3
pir	98.9	1.3	100.0	-0.1	99.3	0.2	97.8	2.3	99.5	0.7
pis	89.0	8.4	93.6	8.0	90.6	12.6	86.7	18.4	91.5	13.7
piu	95.2	-2.7	98.4	-1.6	97.7	-1.4	92.0	14.2	97.9	2.2
pjt	95.4	1.2	98.8	-1.1	98.8	-0.9	94.1	-8.5	98.3	-1.1
pkb	86.6	8.5	94.3	4.6	91.3	2.7	89.5	3.6	91.5	-2.1
pko	92.1	8.0	93.9	9.6	93.1	6.8	93.5	1.3	95.8	2.3
pkr	83.0	-0.1	90.3	-5.4	97.7	0.7	93.7	0.1	96.0	-2.6
plg	98.8	-1.6	99.1	-1.4	96.4	-5.7	97.4	-1.8	98.2	-2.6
pli	57.2	4.7	54.1	60.3	48.5	68.0	68.1	48.3	34.7	78.2
pls	98.9	0.9	99.3	0.3	98.8	-0.3	98.5	0.6	99.7	0.3
plt	99.3	-1.1	99.2	-1.6	99.5	-0.4	99.5	-0.5	99.8	0.1
plu	97.9	1.7	98.7	-2.4	99.0	-0.7	98.7	0.0	99.2	-0.6
plw	98.2	-1.5	97.6	-4.4	99.3	-0.9	98.5	-1.2	99.6	-0.2
pma	97.7	-4.0	98.4	-3.2	98.3	-2.1	98.1	-1.5	98.4	-1.4
pmf	92.4	7.9	97.7	3.1	96.3	4.0	92.6	11.7	95.2	7.6
pmq	83.8	27.6	83.3	23.5	82.7	25.3	83.2	25.6	83.4	25.3
pms	95.7	2.0	96.9	-3.5	96.7	-1.2	93.1	5.3	98.3	2.2
pmx	95.7	4.1	96.2	2.0	94.2	-0.4	92.6	-1.5	95.7	1.0
pnb	70.9	-3.5	71.6	-13.4	92.0	4.2	71.9	-15.5	82.5	23.2
pne	42.5	24.5	7.6	74.4	54.2	5.1	59.5	-20.6	50.8	17.5
pny	98.4	2.2	98.8	2.1	97.5	-0.6	98.3	1.0	98.4	1.2
pnz	96.3	5.0	96.9	3.6	93.1	-1.7	95.2	2.5	95.7	1.1
poe	99.0	-0.2	98.5	-2.3	99.5	-0.0	99.3	1.0	99.6	0.2
poh	99.6	-0.5	99.9	-0.1	99.6	0.0	99.4	0.7	99.7	0.0
poi	98.4	-1.0	99.4	-0.9	99.0	0.4	99.1	0.7	99.5	0.0
pol	94.4	-5.1	93.9	-8.0	94.9	-6.0	88.8	-16.7	97.0	-3.8
pon	93.0	-1.5	97.3	1.8	95.7	2.4	94.1	2.8	96.7	3.4
por	84.9	-3.5	90.6	-4.9	90.3	-7.0	79.0	-19.1	93.5	-5.4
pos	97.2	3.4	99.0	1.6	98.0	1.9	97.6	0.3	98.9	1.0
pot	99.4	-1.3	98.7	-2.6	99.8	-0.3	100.0	0.0	100.0	-0.1
pov	52.9	-32.3	67.4	-27.7	73.8	-17.6	72.8	-15.2	82.2	-13.3
poy	93.2	-3.4	98.1	-2.1	96.6	-2.5	93.9	-1.3	97.0	-1.3
ppk	93.7	3.9	95.4	3.4	94.1	2.3	91.7	-0.5	95.2	1.1
ppo	98.8	-2.1	98.8	-2.2	99.7	-0.1	99.8	0.4	99.7	0.1
pps	97.9	1.1	99.2	0.2	97.9	0.2	96.2	-2.2	98.5	0.2
prf	93.9	0.2	98.5	1.1	97.2	-0.6	95.0	-1.9	98.1	-1.6
pri	95.5	4.0	97.8	3.2	92.7	-1.1	93.4	-4.3	94.4	-4.2
prk	67.2	-2.5	74.9	-39.5	63.9	-0.2	67.4	-6.0	60.0	1.8
prq	87.0	-2.8	95.0	-2.1	95.9	-1.5	92.7	-1.0	95.4	-3.8
prs	50.2	-26.5	44.7	-59.6	76.8	-10.2	54.8	-42.7	76.2	-18.2
prt	94.0	1.0	97.2	3.7	98.0	2.8	97.6	0.8	98.4	1.9
pse	91.8	4.9	95.5	3.3	93.9	0.8	92.8	0.6	93.7	-0.5
pss	98.0	-3.3	99.8	-0.2	99.5	-0.6	98.8	1.3	99.5	0.0
pst	60.5	-1.1	71.3	17.1	61.4	-39.7	49.4	-15.7	74.0	-20.2
ptp	99.5	0.4	99.7	0.1	99.6	0.4	99.5	0.7	99.4	0.2
ptu	84.8	9.7	93.9	2.5	94.9	4.3	89.8	6.5	95.0	6.9
pua	91.5	-1.7	93.6	-0.8	90.9	6.0	88.6	8.9	91.9	7.3
pui	98.0	-3.5	98.8	-2.4	99.8	-0.3	99.9	-0.1	99.9	-0.1
puu	90.6	-0.8	94.6	0.6	92.5	-0.2	91.3	0.1	94.7	1.0
pwg	88.7	-3.5	98.9	-0.1	97.2	1.2	95.0	1.0	97.2	0.1
pwo	99.5	0.5	93.5	-11.8	98.5	2.9	99.2	1.5	99.7	0.5
pww	95.7	1.6	97.1	4.2	97.4	0.6	98.9	-1.0	99.5	0.1
pxm	99.4	1.1	99.3	1.2	98.6	1.5	99.0	0.2	99.7	-0.2
qub	76.6	-11.2	89.6	-1.1	87.0	2.3	80.3	3.1	82.4	12.8
quc	95.0	-3.1	99.2	-0.2	97.9	-0.8	97.0	2.1	99.0	0.5
quf	88.2	-4.9	93.1	-8.4	94.7	-1.5	89.3	5.3	95.7	-0.2
qug	66.8	2.8	89.8	9.0	88.2	4.2	84.1	-4.0	90.5	-0.2
quh	80.7	-2.1	89.9	-6.4	90.2	4.0	83.3	8.9	90.7	6.1
qul	77.8	3.6	89.6	5.3	88.0	-4.5	80.7	-6.8	88.6	-7.2
qup	74.9	-14.9	78.4	-33.6	84.5	-17.6	70.0	-32.5	80.6	-21.9
quw	85.1	2.4	96.3	1.8	93.4	-6.2	88.7	-1.8	96.4	-2.5
quy	98.3	2.8	98.8	2.1	98.2	1.9	98.6	1.6	99.0	1.5
quz	80.3	1.0	92.1	7.4	91.5	1.7	86.8	0.7	94.0	3.3
qvc	94.2	-3.7	98.2	-3.5	97.5	-2.6	94.2	-7.4	97.5	-2.5
qve	78.7	4.9	92.7	-4.6	93.0	-2.9	89.1	-0.4	94.3	-0.4
qvh	73.6	14.3	89.9	4.7	89.6	2.2	84.2	-8.0	91.1	6.1
qvi	71.1	-5.5	90.3	-4.5	90.9	-2.4	84.7	7.4	92.8	1.9
qvm	75.1	18.5	91.2	2.8	90.1	-1.1	82.0	14.7	91.3	-0.6
qvn	89.6	-3.3	97.2	-1.0	96.4	-0.4	90.4	0.3	96.0	-0.9
qvo	85.6	1.0	93.5	-10.6	95.6	1.1	92.8	-1.0	93.2	0.4
qvs	65.7	8.4	59.6	36.3	77.1	13.2	43.5	30.8	70.7	21.8
qvw	95.5	-2.1	96.5	-1.3	95.9	-0.2	95.5	-2.7	97.7	-0.4
qvz	87.1	-5.4	96.0	-2.6	93.5	1.0	88.6	1.3	94.9	4.6
qwh	94.5	0.1	97.3	-3.4	97.3	-2.0	96.6	2.1	98.0	-0.2
qws	53.7	41.9	66.3	47.5	73.7	36.4	74.3	30.9	76.1	36.5
qxh	77.2	-4.4	93.6	-2.6	92.0	-2.4	86.6	-4.9	87.6	-12.6
qxl	96.5	-5.0	98.8	-1.9	99.3	-0.8	99.3	-0.4	99.6	-0.3
qxn	42.5	-16.0	76.6	-18.8	83.7	-6.1	71.3	-27.6	84.9	-10.4
qxo	62.7	-8.6	82.8	8.7	83.6	3.2	65.9	31.6	84.0	7.1
qxr	72.4	2.4	89.3	9.8	87.6	-2.9	83.6	-6.1	86.9	-1.7
rad	82.8	-24.7	95.5	-3.9	98.2	1.9	97.2	0.7	98.5	1.8
rag	81.2	7.5	94.2	4.9	91.9	6.1	85.2	7.7	91.2	11.1
rah	92.4	0.4	96.7	-0.6	96.9	-2.3	95.1	-7.1	97.8	-2.6
rai	96.5	-1.2	98.9	-0.0	97.6	1.4	96.9	2.0	98.1	1.0
ram	97.8	2.8	97.9	3.5	99.5	0.1	99.3	0.1	99.3	0.2
rap	99.2	-1.3	99.4	-1.3	99.8	-0.1	99.8	0.1	99.9	0.0
rar	60.3	1.9	76.2	17.7	78.8	-14.4	73.6	-13.0	84.6	-7.3
rav	94.0	-5.2	98.4	-3.1	99.8	-0.0	98.5	-2.8	99.8	-0.2
raw	99.8	-0.4	98.3	-3.2	100.0	-0.1	100.0	0.0	100.0	-0.0
reg	95.0	-0.5	98.8	-1.9	96.2	-4.4	92.4	-7.7	94.9	-6.7
rej	55.1	-35.6	71.5	-12.4	81.7	12.8	74.6	3.0	84.0	12.1
rel	93.1	9.3	97.5	3.0	94.1	0.5	88.5	-2.2	92.5	-3.0
rgu	78.6	0.2	92.4	1.9	88.2	5.9	82.4	3.6	89.4	7.1
rhg	98.6	-2.3	99.1	-1.9	99.5	-0.9	99.7	0.0	99.7	-0.4
ria	95.1	5.6	96.0	3.7	95.9	4.2	94.6	2.5	97.9	1.3
rif	64.3	52.0	69.6	45.5	95.8	-2.6	96.7	0.6	98.0	0.1
rim	93.5	7.8	95.7	5.5	92.2	1.3	90.1	-4.1	92.5	-3.8
rjs	91.1	-7.1	98.1	-3.7	99.4	-0.3	98.1	-2.7	99.7	-0.0
rkb	92.0	-11.6	99.2	-0.9	98.3	0.8	97.5	0.3	98.7	1.0
rkt	83.5	-8.9	87.6	-20.6	98.8	1.7	96.3	2.6	99.5	0.7
rmc	91.5	2.5	93.7	-3.4	95.2	4.6	94.4	4.7	96.1	4.0
rml	95.4	3.6	93.4	-8.6	98.2	2.7	97.6	-0.6	98.8	1.8
rmn	80.9	-28.7	82.1	-28.6	98.2	0.1	93.6	-6.8	96.4	-3.7
rmo	91.2	1.5	95.8	2.3	91.2	-5.0	87.9	-6.2	93.1	-4.0
rmy	16.9	57.4	32.9	68.3	78.9	-7.2	73.2	1.6	86.5	-3.5
rmz	95.0	4.6	96.4	6.1	98.4	0.0	97.8	1.4	98.5	0.9
rng	78.4	16.0	91.5	8.5	85.9	0.6	79.9	-5.1	86.4	-6.6
rnl	46.7	0.6	63.5	51.9	67.7	16.5	53.8	11.9	58.9	16.2
roh	67.2	-26.0	84.5	-13.0	81.2	-12.3	59.7	30.3	84.0	5.2
ron	77.2	30.1	87.8	14.5	92.5	-4.4	90.9	-1.2	95.7	-3.8
roo	99.1	-1.4	98.4	-3.1	99.3	-0.4	99.6	0.4	99.8	-0.1
rop	97.6	-1.6	99.1	-0.4	98.5	0.6	96.9	4.7	97.7	-0.2
row	77.9	11.7	92.8	3.0	86.8	8.1	76.7	23.6	89.3	3.3
rro	94.5	5.3	98.3	1.2	95.5	-0.3	92.2	0.7	95.1	-2.2
rub	92.5	12.4	96.8	2.1	94.2	0.7	91.8	-0.8	94.9	-0.4
rue	54.3	-9.8	67.1	-34.7	91.2	3.3	67.7	39.0	96.0	5.7
ruf	89.2	8.1	97.8	-0.4	94.6	-1.5	92.1	-0.2	95.6	-1.4
rug	88.9	-0.5	96.9	0.4	94.7	7.5	91.9	11.9	94.0	10.9
run	61.5	-0.3	76.6	4.1	67.5	-0.9	53.0	7.7	75.4	-0.0
rup	84.2	10.8	89.6	5.8	84.4	-0.1	81.3	1.5	90.8	0.1
rus	51.6	-41.3	58.0	-51.1	73.4	-18.5	62.3	-29.0	79.8	-14.2
rwk	86.4	-16.8	96.1	-6.8	95.2	-6.0	93.6	-1.8	95.0	-6.8
sab	98.2	-1.7	99.6	-0.3	98.4	-1.9	98.8	-0.8	99.5	-0.3
sac	88.7	20.2	88.6	20.3	87.7	19.7	88.3	20.3	88.5	18.9
sag	91.6	7.5	96.5	5.2	93.9	3.9	93.3	0.6	97.4	1.4
sah	86.3	3.5	90.5	10.0	91.6	-3.1	87.5	-1.9	95.4	1.6
saj	83.9	8.9	90.6	13.5	90.2	6.3	88.3	3.4	91.6	-1.6
san	58.6	53.6	85.1	18.7	88.8	-10.2	87.1	-10.0	82.6	-22.5
saq	72.5	21.1	87.0	12.2	84.5	-6.7	77.4	-10.4	83.6	-11.5

Table 8: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
sas	80.6	27.3	82.1	28.4	74.7	3.9	73.2	4.8	69.8	-9.9
sat	66.0	49.4	62.8	53.3	95.1	2.6	95.5	0.6	96.7	0.3
saw	70.6	14.4	76.2	26.8	75.6	5.6	64.8	-18.3	79.9	-0.1
saz	93.5	4.3	82.8	24.7	96.5	-4.3	97.4	-2.5	98.1	-2.2
sba	97.7	3.5	98.0	2.4	96.6	4.4	96.5	4.4	97.4	4.9
sbd	89.2	7.8	93.7	9.2	91.0	5.1	89.4	-1.2	95.5	1.4
sbe	93.1	5.1	98.7	-0.7	95.8	-2.1	94.3	0.1	96.3	-2.3
sbk	88.4	-0.4	96.4	-4.6	92.1	1.1	86.8	13.1	91.9	6.1
sbl	89.0	-10.2	97.9	-1.4	97.7	0.7	97.4	1.9	98.9	0.6
sbp	79.0	12.6	92.9	7.6	90.3	3.8	86.2	4.4	90.9	7.2
sbs	79.9	-19.1	97.0	-3.2	94.7	-4.2	88.2	-2.3	94.4	-3.4
sby	81.2	-2.3	94.7	-6.9	93.3	-1.7	84.8	-5.1	93.2	-4.3
sch	93.7	-6.0	95.3	-8.5	97.5	-1.8	96.9	3.5	94.3	-6.8
sck	78.7	-7.6	96.4	-3.1	97.1	-1.6	93.5	-7.4	98.6	1.1
scn	88.5	-13.3	87.7	-18.3	93.0	-6.5	91.8	5.9	95.9	-0.6
sco	57.0	-34.5	68.7	-40.1	86.8	7.5	55.2	47.8	93.2	10.7
scp	83.7	12.3	87.4	9.6	86.0	1.8	85.5	7.7	86.7	2.4
sda	89.8	-2.1	95.3	4.1	94.2	-2.1	94.2	0.4	94.8	-2.2
sea	96.1	-4.2	96.8	-5.5	98.1	-1.3	98.6	0.3	99.1	1.1
see	100.0	0.0	98.7	-2.5	100.0	-0.1	100.0	0.0	100.0	0.0
sef	89.1	8.9	92.3	10.6	90.9	4.3	87.6	-5.9	91.6	-0.9
seh	74.1	-0.7	93.9	-5.5	87.7	-10.6	79.2	-20.1	87.1	-15.2
sei	69.7	-17.4	85.5	23.6	85.0	23.8	84.9	23.0	84.5	21.4
sev	80.5	27.0	87.5	21.1	87.4	10.3	83.8	6.5	88.0	8.0
sey	99.1	1.8	99.8	-0.1	99.3	1.0	98.9	0.9	99.3	1.2
sfw	88.4	3.9	93.8	7.1	88.7	-1.2	83.7	-4.3	90.6	-1.4
sgb	95.3	-1.6	98.9	-0.9	98.1	-1.7	96.9	-1.1	98.5	-0.1
sgs	89.4	-11.1	74.5	-33.7	92.5	-1.6	78.4	33.1	93.3	11.2
sgw	92.3	2.5	87.1	19.0	97.7	-1.0	97.8	1.4	98.6	0.1
sgz	95.2	-2.5	99.6	0.0	97.5	2.1	97.2	4.0	98.3	1.8
shb	99.5	-1.0	99.7	-0.7	99.7	-0.6	100.0	0.0	99.9	-0.1
shi	82.7	26.9	93.7	5.9	90.8	-0.3	90.3	1.2	93.4	-0.9
shk	97.3	4.0	97.2	2.6	96.5	3.3	97.9	0.7	98.3	2.6
shn	97.7	-1.5	99.9	0.2	99.9	0.0	99.9	-0.2	99.9	0.1
sho	95.9	-6.6	99.1	-1.4	98.8	-1.2	96.4	-5.4	99.2	-1.2
shp	96.4	-2.3	99.0	-1.0	96.9	0.0	95.6	6.2	97.5	2.6
shu	85.0	-7.9	76.2	-26.1	96.2	4.2	95.8	5.2	98.4	1.3
sid	84.4	-7.0	92.8	7.4	91.2	-6.7	86.9	-11.7	94.8	-4.3
sim	96.8	-5.2	99.1	-1.4	99.0	-1.8	99.4	0.5	99.6	-0.1
sin	99.9	0.0	99.2	1.0	99.2	0.1	99.9	0.0	99.9	-0.2
sja	96.6	3.0	98.5	-1.3	97.9	1.0	96.2	1.6	98.6	0.9
sjm	89.5	8.7	54.1	-51.2	92.9	9.3	90.8	7.2	94.2	7.2
skn	50.9	-48.6	94.9	6.2	92.7	2.7	90.0	9.1	95.7	2.1
skr	69.7	-10.2	65.4	-28.8	88.7	0.4	39.7	56.6	84.1	-13.7
sld	99.1	-1.1	99.7	-0.7	99.9	-0.2	99.5	0.2	99.1	-1.2
sli	65.8	-2.3	77.3	-10.1	82.2	-1.9	67.8	-3.3	75.9	-23.1
slk	87.0	2.8	90.1	6.8	87.1	-1.2	84.0	-15.8	94.9	-1.1
sll	98.6	-0.8	99.3	-1.1	98.9	-0.6	97.8	-2.0	98.7	-1.2
slv	77.4	-2.4	89.8	12.4	87.7	-4.1	77.3	-15.5	92.8	-3.8
sme	73.4	-32.0	89.8	-8.1	87.4	-11.1	88.4	11.6	96.5	3.3
smk	93.6	-1.0	98.7	0.1	96.4	-0.2	94.7	4.2	97.3	1.1
smo	86.4	7.9	87.1	3.5	87.8	-10.4	86.5	-6.3	93.6	-2.5
smt	85.9	-2.2	91.0	-13.6	91.9	-8.2	87.5	-7.2	93.3	-7.3
sna	79.1	-6.2	94.7	-1.0	92.3	-5.5	85.1	1.0	94.3	0.6
snc	98.1	0.6	97.8	-3.3	97.9	-0.8	98.3	1.9	97.9	-0.8
snd	90.4	13.7	90.4	12.8	93.0	-1.3	92.7	-0.2	97.5	0.7
sne	99.1	-1.3	99.4	-1.2	99.4	-1.0	99.9	-0.2	99.8	-0.3
snn	99.4	0.2	99.5	-0.8	98.4	-2.1	96.1	-6.0	99.1	-0.7
sny	98.7	-1.9	99.5	-0.3	99.2	0.8	99.3	1.0	99.4	0.5
som	86.2	-18.2	94.9	-6.9	88.8	-13.7	75.8	-28.4	90.5	-10.5
sop	91.8	3.4	94.0	-1.2	95.1	4.1	92.5	5.3	94.8	6.3
soq	94.0	-8.3	98.8	-2.3	98.2	-2.6	97.7	-0.3	98.4	-2.0
sot	82.0	5.5	91.2	8.6	87.4	10.8	82.9	13.5	91.2	8.1
soy	98.6	-2.2	99.1	-1.9	99.3	-0.9	99.6	-0.4	99.5	-0.7
spa	70.1	-19.7	84.8	-14.2	82.9	-10.7	66.6	-32.4	86.4	-15.2
spl	97.3	0.6	99.5	0.2	98.6	2.4	97.7	2.6	98.8	2.1
spp	98.8	0.0	99.3	-0.4	98.7	0.9	97.8	2.4	99.0	0.1
sps	98.6	-2.5	99.6	-0.8	99.3	-1.0	99.3	0.8	99.6	-0.4
spy	98.5	1.0	99.0	-0.8	98.4	-0.1	98.5	0.7	99.3	0.5
sri	98.4	1.6	99.2	0.4	98.6	1.0	98.3	1.8	98.7	1.3
srm	92.7	3.6	97.8	4.1	95.8	6.1	93.8	8.6	96.5	5.3
srn	89.3	-1.5	96.9	1.3	94.0	-1.9	90.9	-2.6	95.6	-2.5
srp	0.3	2.8	1.2	10.0	63.7	-3.4	51.7	-27.3	70.7	-15.1
srq	96.1	-1.7	98.8	0.8	97.8	1.6	97.6	1.7	98.0	0.9
srx	84.7	7.1	97.2	-3.9	97.3	0.8	94.8	-1.1	98.2	0.2
ssc	89.2	8.6	94.7	-1.8	92.1	-2.9	89.7	3.6	92.6	0.8
ssd	97.7	-2.5	99.2	0.0	98.2	-1.2	98.3	0.5	98.7	-0.5
sse	91.2	-11.6	98.9	-2.2	98.3	-2.7	98.1	0.2	99.2	-1.4
ssg	95.3	2.4	98.1	-0.3	95.3	-2.8	94.1	0.4	96.8	0.3
sso	93.0	8.5	96.2	3.6	95.0	0.9	88.5	-7.7	94.8	-0.4
ssw	73.7	16.7	89.1	13.1	81.9	0.0	75.2	1.8	89.0	5.9
ssx	99.9	-0.2	97.2	-5.4	99.8	-0.1	99.7	0.1	99.9	0.1
ssy	66.1	-18.2	88.6	15.5	83.5	7.4	80.5	3.8	83.9	5.6
ssz	95.1	1.7	98.5	1.8	96.5	-1.5	95.2	-2.2	96.8	-1.6
stn	98.4	-2.7	99.2	-1.7	99.1	-1.3	99.6	-0.2	99.6	-0.7
stp	99.9	-0.1	99.8	-0.5	99.7	0.0	99.8	0.2	99.8	-0.2
stq	62.4	-37.0	78.7	-26.0	91.3	4.0	41.8	70.3	81.7	29.2
sua	97.7	-3.4	99.4	-0.5	99.4	-0.4	98.9	-0.5	99.5	0.1
suc	94.1	-6.3	96.3	-6.1	95.9	-1.6	93.9	-5.5	98.0	0.1
sue	97.1	-4.1	99.2	-1.1	99.2	-0.3	97.0	-2.7	99.4	-0.4
suk	95.7	6.7	96.6	4.9	94.3	2.3	94.8	-2.3	95.8	-0.8
sun	68.4	-23.8	88.4	-4.6	85.6	-11.0	78.1	-15.1	94.0	4.1
suo	96.5	3.2	96.9	-0.3	95.2	-1.9	94.1	-3.0	96.8	-1.4
suq	94.0	-3.0	98.3	0.1	97.5	-0.6	96.2	-1.7	97.3	-1.3
sur	99.6	-0.6	99.8	-0.2	99.8	-0.2	99.8	0.4	99.9	-0.0
sus	64.6	51.2	51.5	64.7	95.9	2.5	94.2	2.4	96.0	0.9
suz	90.9	-7.1	98.4	-1.8	99.1	0.1	97.7	-2.3	99.6	0.5
swb	86.9	12.3	92.8	10.5	84.0	-8.9	83.1	-6.5	86.9	-9.9
swc	42.4	9.4	72.1	29.8	71.6	19.5	61.0	4.9	67.9	26.9
swe	90.1	6.1	93.1	3.5	87.6	-0.6	75.5	-21.1	94.3	2.3
swh	52.1	-44.3	77.8	-28.6	78.7	-18.1	68.6	-2.8	79.3	-18.9
swk	70.5	15.8	87.1	19.8	83.7	8.7	76.0	0.6	82.1	15.3
swp	95.0	-4.3	98.4	-2.2	97.1	-0.8	96.3	0.8	97.3	-1.7
swv	62.6	22.6	87.5	14.1	88.0	4.4	83.9	4.2	89.1	2.7
sxb	94.7	-1.4	98.5	-2.9	97.2	-3.6	95.5	-1.5	96.6	-4.0
sxn	84.8	-11.7	85.5	24.0	83.7	5.6	79.5	-5.1	84.3	11.5
syb	97.0	-3.1	99.5	-0.6	98.9	-1.7	98.8	-1.0	99.6	-0.1
syw	94.3	6.2	95.9	7.3	95.4	1.7	96.0	3.9	95.5	-0.9
szb	68.0	-15.7	62.4	-29.4	79.5	2.5	80.0	6.9	84.2	12.6
szl	85.0	-15.1	84.4	-19.8	92.0	-7.3	93.3	5.2	98.3	0.3
szv	97.7	2.4	98.4	2.5	98.1	0.3	98.0	0.1	98.0	-0.1
szy	22.1	55.5	64.5	36.2	73.5	5.3	77.5	3.2	86.7	-5.4
tab	87.8	3.3	93.9	-3.5	95.0	1.3	90.1	3.6	95.9	2.5
tac	98.0	0.1	97.6	-3.5	98.7	-0.4	98.6	0.7	99.2	-0.5
taj	93.7	-1.7	96.5	-5.9	99.0	0.2	97.1	3.2	99.2	0.5
tam	92.7	-4.9	86.0	-21.1	98.2	1.0	97.9	1.0	98.8	1.0
tan	97.3	0.4	98.7	-0.9	95.5	-2.2	97.3	-1.7	98.4	0.5
tao	92.1	-5.2	97.7	0.5	96.9	-1.3	93.7	-3.1	96.7	-1.8
tap	87.8	-14.1	96.5	-6.5	98.3	-0.6	96.8	0.2	98.7	-0.8
tar	98.1	-3.6	99.7	-0.5	99.7	-0.4	99.7	-0.4	99.8	-0.0
tat	75.4	-5.6	83.6	-7.5	89.6	-1.9	83.4	-13.1	94.5	1.4
tau	99.0	1.1	97.6	-1.7	98.8	0.7	99.2	1.4	99.2	1.5
tav	96.1	-5.3	99.4	0.5	98.6	1.3	96.9	-1.2	98.9	0.1
taw	96.5	0.7	99.4	-0.8	99.6	-0.5	96.4	-1.8	99.3	0.8
tay	94.2	-7.1	97.1	-5.2	97.2	-4.2	96.4	-5.5	97.2	-4.3
tbc	98.3	-1.0	99.4	-0.9	98.8	0.1	98.6	0.4	99.1	0.2
tbf	94.0	-2.6	97.5	1.6	96.5	2.9	95.0	3.3	96.7	1.4
tbg	86.9	-23.0	88.4	-20.9	98.7	-1.8	98.7	-1.8	98.5	-2.3
tbk	95.2	2.4	97.7	0.0	97.8	-1.2	96.2	-1.7	98.6	-0.1
tbl	96.4	3.9	97.4	4.5	97.1	2.8	96.7	4.7	97.2	2.8
tbo	95.6	0.1	98.4	0.9	96.1	1.9	95.8	2.4	96.9	0.5
tbw	87.5	18.1	90.8	11.6	90.8	9.8	89.3	7.8	91.8	10.0
tby	93.8	-10.5	97.6	-4.5	98.7	-1.8	99.0	0.3	99.8	0.1
tbz	98.4	-2.0	97.4	-5.1	99.0	-0.6	98.5	-2.1	99.7	0.2
tca	99.6	0.6	99.3	-0.1	99.0	1.9	99.0	1.9	99.1	1.8
tcc	97.3	1.9	97.6	1.9	98.0	0.9	97.9	-0.4	98.7	-0.7
tcd	86.1	3.1	95.6	1.5	92.9	2.3	88.9	7.2	92.9	-3.0
tcs	94.3	-5.0	98.8	-0.3	95.6	4.6	94.4	4.0	97.1	0.4
tcy	76.3	-21.4	84.1	-0.1	77.3	-33.9	80.3	-22.3	96.1	-3.7
tcz	86.5	7.1	90.0	8.4	90.5	4.3	86.7	4.3	90.5	0.5
tdc	92.1	4.4	94.3	5.5	93.5	-3.7	91.2	-5.5	93.9	-4.9
tdg	91.0	3.0	96.2	2.1	96.1	0.9	93.0	-0.2	94.3	1.7
tdh	83.0	12.1	79.8	-7.3	83.8	-1.3	81.0	-4.6	84.7	2.1
tdt	86.2	-10.3	95.0	-3.4	92.1	-3.5	87.1	-9.2	95.3	-4.0
tdx	98.4	0.5	99.4	-0.5	98.4	-0.3	98.5	0.8	98.6	0.2
ted	94.5	3.0	98.2	1.9	97.3	1.5	95.1	1.3	96.9	-1.4
tee	98.9	-1.1	99.7	-0.4	99.4	-0.1	99.4	0.1	99.6	0.0
tel	85.7	6.4	86.4	13.7	98.2	-0.9	98.4	0.7	99.1	0.7
tem	97.2	2.9	99.4	-0.8	94.0	8.1	95.0	8.7	96.2	5.9
teo	29.2	-71.0	40.4	-71.5	88.5	-12.9	86.3	-11.1	92.2	-8.0
ter	98.3	-3.3	99.1	-1.8	98.7	-1.7	99.5	-0.6	99.4	-0.1
tet	76.9	-0.6	90.3	-2.1	87.5	4.3	80.5	4.8	92.1	6.3
tew	99.8	-0.3	99.3	-1.2	99.5	-0.9	99.9	0.1	99.5	-0.8
tex	94.9	5.4	96.5	5.4	94.7	0.5	93.2	-1.8	93.9	-3.3
tfr	98.3	2.3	99.2	0.7	97.0	2.1	97.8	1.5	98.5	0.8
tgj	88.3	13.5	86.1	6.3	88.9	-0.9	89.4	1.4	88.6	-3.6
tgk	79.5	-6.3	77.9	-20.9	90.6	-5.5	88.1	0.4	96.5	-1.8
tgl	58.3	-23.0	76.8	-6.8	69.3	17.9	57.1	30.7	65.1	41.0
tgo	99.1	0.5	99.8	0.1	99.0	0.2	98.9	0.7	99.1	0.3
tgp	94.7	-2.7	98.9	0.9	97.3	-1.4	93.8	-2.6	97.1	-4.5
tgw	91.6	7.6	95.5	5.3	94.4	4.5	93.6	4.7	95.2	5.5
tha	95.7	-6.9	97.2	-4.4	98.1	-2.8	99.0	-0.5	99.7	0.1
thk	78.3	10.1	89.6	4.9	86.5	-8.0	78.2	2.9	85.0	-14.1
thl	73.4	10.7	84.3	0.7	85.2	-3.0	81.5	6.6	84.1	-7.5
thq	72.6	12.0	83.0	0.0	92.9	-0.6	85.8	-14.7	94.3	-0.9
thr	76.1	-4.4	96.5	-4.2	97.6	-1.8	96.6	0.1	96.7	-4.5
tif	97.3	-5.0	99.0	-1.5	99.6	-0.7	97.9	-3.7	99.3	-0.8
tih	78.7	19.6	84.2	24.7	70.2	-15.3	69.3	-15.5	83.1	20.1
tik	99.6	-0.9	98.7	-2.5	99.7	-0.7	99.9	-0.3	99.8	-0.5
tim	98.9	-1.4	99.9	-0.2	99.7	-0.6	99.1	-1.0	99.9	-0.1
tir	94.3	1.3	85.2	-20.2	95.5	-3.0	94.0	-5.6	96.8	-2.7
tiu	92.5	9.7	96.0	6.8	93.2	1.9	92.0	1.4	93.7	-2.6
tiv	91.6	7.3	95.7	4.7	94.9	3.4	93.9	2.7	96.9	4.2
tiy	95.7	-7.2	97.6	-4.0	99.2	-0.4	99.4	0.7	99.4	0.2
tke	88.6	-1.7	96.5	-4.0	96.1	0.4	93.8	-0.2	97.5	1.7
tkl	82.6	12.0	87.8	8.2	85.2	-2.7	83.6	-8.5	89.3	-2.9
tkr	96.1	-5.4	98.9	-1.0	98.1	-1.5	98.1	-0.1	98.9	-0.9
tku	93.3	-3.1	94.6	0.2	96.7	-2.1	94.3	-0.5	96.9	-1.8
tlb	88.9	9.1	93.9	8.9	86.8	-2.7	80.5	-13.5	83.0	-14.1
tld	80.5	30.8	84.0	24.1	87.0	-1.8	84.7	1.2	88.8	-12.0
tlf	97.2	0.2	99.7	-0.4	99.1	-0.6	97.7	2.7	99.4	-0.7
tlh	97.2	-1.9	97.8	2.2	97.9	-0.4	96.9	-4.1	98.6	-0.5
tll	86.8	7.3	92.3	8.2	89.1	6.2	86.2	10.5	90.1	6.2
tlm	92.5	-5.8	96.0	-1.5	97.2	-3.8	94.5	-7.3	97.0	-4.8
tly	95.5	5.6	95.3	-1.2	95.6	1.4	95.1	2.0	97.9	0.6
tmc	100.0	0.0	99.7	-0.7	99.9	-0.0	100.0	0.0	100.0	0.0
tmd	98.7	1.2	99.4	0.6	98.0	1.7	97.0	0.9	98.8	1.4
tmf	99.2	-1.3	99.0	-1.9	99.7	-0.4	99.9	0.2	99.7	-0.3
tml	76.6	30.9	79.5	31.5	68.7	-11.6	65.9	-12.4	74.0	7.3
tna	95.3	-4.7	97.9	-2.4	96.8	-3.2	97.4	-2.0	98.2	-1.2
tnc	98.9	0.9	98.1	-1.5	98.6	0.4	99.2	1.1	99.3	1.3
tng	99.3	-1.1	99.7	-0.6	99.5	-0.8	99.7	-0.2	99.9	-0.1
tnk	98.8	1.6	99.5	0.3	98.8	0.1	98.3	0.8	99.1	1.3
tnn	96.4	-2.4	99.0	-1.1	97.7	0.3	96.0	0.9	98.5	0.5
tnp	97.8	0.6	99.2	0.1	98.1	1.5	96.5	1.7	98.9	0.9
tnr	98.5	0.8	98.9	-1.3	98.5	0.4	97.6	0.0	99.3	0.3
tnt	74.8	-5.3	86.2	10.5	81.8	5.4	85.0	17.1	84.7	5.2
tob	94.8	-6.7	98.8	0.2	96.6	0.4	95.7	-2.6	98.3	0.8
toc	99.5	1.0	99.7	-0.5	99.5	0.4	99.6	0.8	99.7	0.7
tod	97.6	2.5	98.4	-2.1	96.4	1.2	97.6	3.1	96.2	6.2
tof	97.3	-0.5	99.0	-0.3	97.5	1.6	97.8	0.0	98.5	1.5
tog	86.5	-2.0	90.1	-6.9	93.2	-1.1	87.0	-2.3	95.2	0.7
toh	89.3	9.9	94.5	7.3	92.6	2.0	87.9	-3.3	94.4	2.7
toi	67.4	-7.8	86.5	5.1	85.1	5.2	74.1	-2.2	77.5	17.4
toj	96.6	-3.6	99.1	0.7	98.8	0.9	98.5	0.6	98.8	0.7
tom	84.2	6.8	89.5	12.8	86.0	8.8	85.0	9.7	80.3	-6.1
ton	96.3	2.1	96.7	-1.4	95.3	0.4	96.4	5.3	98.1	1.3
too	93.7	4.2	95.0	0.3	96.7	1.7	94.6	-0.2	96.9	2.1
top	96.7	-2.2	98.9	0.3	98.9	-0.5	96.9	3.6	99.1	-0.8
tos	97.1	2.1	99.2	-0.5	99.2	-0.2	97.2	-3.8	99.6	0.2
tpi	85.0	-0.2	91.1	5.7	87.4	-1.7	78.8	-12.1	90.9	-0.7
tpm	94.2	2.5	98.2	2.5	96.0	0.5	93.9	3.5	97.6	-0.3
tpt	98.6	-0.5	99.8	-0.2	99.2	-0.3	98.9	-0.5	99.2	-0.4
tpz	99.2	1.1	99.8	-0.0	98.2	-0.7	98.4	-0.6	98.9	-0.1
tqo	78.4	4.6	81.8	5.6	86.5	3.9	83.7	-0.9	87.6	6.4
trc	99.4	-0.9	99.6	-0.8	99.8	0.2	99.8	0.1	99.9	0.1
tri	99.7	-0.5	99.3	-1.4	99.9	-0.2	99.8	0.0	99.9	-0.2
trn	98.8	-2.2	99.3	-1.3	99.9	-0.3	99.8	-0.2	99.9	-0.2
tro	47.6	-51.2	77.1	-25.9	88.7	-3.3	85.1	-2.7	89.6	-2.9
trp	92.2	8.5	94.3	5.8	93.0	1.8	90.2	0.8	93.8	-0.4
trq	97.7	1.0	96.5	-1.6	98.5	2.6	98.6	2.7	98.6	2.7
trs	99.9	0.2	99.8	-0.1	99.9	0.1	99.9	0.2	99.7	-0.6
trv	90.1	-4.1	97.6	1.4	95.0	-0.4	93.0	0.7	96.8	-2.3
tsa	83.0	-3.9	94.4	5.4	92.6	1.7	86.3	0.3	91.5	0.6
tsc	82.1	5.7	90.2	12.4	86.2	-6.0	84.0	-5.4	89.6	-8.0
tsn	83.4	-7.3	90.9	-4.4	90.8	-6.1	88.1	-7.4	93.8	-3.6
tso	82.2	-12.2	94.6	-4.0	91.6	-3.2	86.1	-2.0	94.5	0.7
tsu	68.4	44.3	71.3	44.1	83.3	13.2	80.6	-11.2	87.3	5.5
tsw	95.2	1.9	99.2	-0.4	96.6	0.1	94.4	1.9	97.7	1.2
tsz	92.4	5.9	95.4	2.2	90.7	-6.6	89.5	-3.9	93.3	-3.6
ttc	99.0	1.2	99.8	-0.2	99.5	-0.6	99.2	-0.0	99.7	-0.4
tte	92.7	-11.6	97.3	-5.1	97.9	-3.6	97.6	-0.8	98.3	-2.0
ttr	89.4	17.4	90.5	11.3	87.6	12.0	82.6	20.3	87.4	8.6
tue	96.8	3.3	99.6	-0.1	98.6	-0.1	97.9	0.7	99.1	-0.2
tuf	98.3	-2.4	99.2	-1.2	99.1	-0.9	99.2	-0.2	99.5	0.0
tui	96.9	1.8	96.5	6.1	95.3	2.4	95.6	-1.0	98.1	2.5
tuk	43.6	63.7	73.4	39.3	88.0	-2.3	78.9	-16.8	93.4	-0.1
tul	93.9	8.2	96.3	6.6	94.1	4.3	94.2	5.0	94.2	2.7
tum	72.5	26.2	82.2	26.7	84.6	3.5	78.7	-0.1	90.9	4.2
tuo	99.0	-0.2	99.6	-0.8	99.7	-0.1	99.6	0.3	99.8	0.3
tur	50.9	-53.0	71.2	-37.7	87.0	-6.4	78.2	-13.8	91.2	-2.6
tuv	87.7	2.9	93.4	8.8	91.0	3.3	87.2	-3.3	90.8	-3.8
tvk	95.5	-4.3	99.3	-1.2	97.0	-0.3	97.2	1.0	98.1	1.4
tvl	85.8	4.2	91.1	7.3	89.1	-1.2	86.4	-8.7	90.8	-0.2
tvs	72.9	-2.3	86.8	-12.9	86.9	7.1	76.0	17.4	86.6	0.5
tvt	96.5	-4.2	99.4	-1.1	98.5	-1.3	97.9	0.7	98.1	-2.0
tvu	96.4	2.5	98.3	1.1	97.3	1.8	96.2	-1.1	97.3	1.4
tvw	93.6	-10.6	97.7	-4.2	98.8	-2.1	98.0	-1.6	99.2	-1.2
twb	67.8	-27.2	85.7	-9.2	77.4	-20.7	69.5	-21.0	76.4	-22.2
twi	89.2	-7.1	95.8	-1.2	93.2	-4.4	88.3	-7.9	96.7	-0.3
twu	65.5	-23.5	87.2	-0.4	84.7	-5.2	84.0	1.3	87.3	1.1
txa	96.0	-4.7	98.6	-2.1	98.7	-1.5	99.1	-0.1	99.4	-0.6
txq	62.8	16.2	88.6	2.9	84.1	3.6	72.1	-5.1	85.6	2.7
txt	74.5	28.8	79.7	29.1	68.6	-3.4	68.6	3.1	63.8	-31.3
txu	99.4	0.1	99.2	-1.6	99.2	-0.5	99.1	-0.8	99.1	-1.3
tye	96.9	-1.2	98.6	-2.4	98.3	-1.0	96.5	4.6	99.0	-0.4
tyv	74.7	1.9	82.8	-7.8	92.0	-0.2	84.8	-7.7	96.1	3.3
tyz	89.8	7.5	93.7	7.6	95.1	-1.2	93.0	2.5	95.8	-3.6
tzh	94.8	3.1	96.2	4.0	93.7	-2.2	93.4	-1.9	95.6	-1.7
tzj	94.5	5.3	99.3	-1.3	98.6	-1.1	97.0	3.0	99.1	-0.4
tzm	1.3	90.2	64.9	48.6	79.5	14.0	73.0	-1.1	77.5	7.7
tzo	96.6	0.9	98.4	-0.5	98.0	1.0	97.2	1.2	98.7	1.0
ubl	93.2	-0.6	97.6	-3.7	97.8	-0.2	96.4	-1.3	99.0	-1.1
ubr	97.0	-3.2	99.4	-1.0	98.4	-2.2	98.3	1.4	99.2	-0.9
udg	82.2	0.2	88.6	7.7	97.4	-0.3	94.8	-2.8	96.8	1.8
udm	79.9	15.1	81.6	-10.0	92.3	10.3	88.2	15.8	97.2	4.0
udu	99.7	-0.4	99.6	-0.8	99.7	-0.3	99.8	0.0	99.9	-0.1
uhn	83.0	-1.6	90.4	5.3	95.4	-1.2	91.2	-6.8	94.1	-5.7
uig	64.8	51.0	77.1	36.6	94.9	0.2	92.5	-1.2	97.1	0.3
ukr	76.6	-7.0	79.9	20.1	84.3	-9.7	70.8	-31.7	90.6	-9.6
ukw	99.3	-0.8	99.3	-1.5	99.4	-0.5	99.0	-0.9	99.7	-0.1
uli	97.0	-4.6	97.0	-5.3	98.2	-1.5	98.8	0.2	99.4	-0.1
umb	90.3	4.4	96.3	2.1	94.3	3.2	92.7	5.1	95.9	4.4
unx	80.2	-20.7	99.0	-0.9	99.3	-0.1	97.5	0.6	99.2	0.2
upv	97.2	-4.8	98.3	-3.0	98.4	-1.9	98.7	-0.8	99.4	-0.5
ura	97.3	-2.6	98.7	-1.4	98.4	0.1	97.8	-0.5	98.1	-0.9
urb	97.8	-0.7	99.2	0.1	98.9	1.2	98.6	0.4	98.9	-0.2
urd	57.8	29.0	28.5	72.0	87.2	-6.4	68.9	-13.2	89.1	3.5
urh	95.6	3.6	96.8	2.9	96.5	4.6	96.5	3.2	97.3	4.5
uri	99.6	-0.6	99.3	-1.4	99.4	-1.0	99.4	-0.0	99.9	-0.2
urk	99.1	-1.7	99.9	-0.3	100.0	0.0	100.0	0.0	100.0	0.0
urt	98.5	1.7	99.7	0.0	98.9	1.2	98.5	2.8	99.4	0.4
ury	98.7	-0.9	99.5	-1.0	99.3	-0.9	99.5	-0.0	99.9	-0.2
usa	98.4	-2.7	99.2	-1.4	98.5	-0.9	98.7	0.6	99.2	-0.1
usp	94.9	-3.9	97.7	-4.3	99.2	-0.8	97.1	-4.2	99.6	-0.5
uth	85.3	18.9	87.3	15.6	91.7	0.5	88.1	-5.1	92.9	1.8
uvh	97.1	-5.5	98.5	-2.9	99.9	-0.3	99.8	0.2	99.8	-0.2
uvl	93.1	-9.6	98.8	-0.4	97.7	1.4	96.7	1.0	97.9	0.8
uzn	48.3	40.0	57.3	56.7	86.3	-4.4	82.9	-13.5	93.8	-2.9
vag	89.6	12.1	94.1	9.8	87.8	-8.2	86.2	-6.3	90.7	-5.3
vah	70.0	-3.6	91.3	-14.4	95.9	5.2	93.2	3.3	97.6	1.0
vai	100.0	0.0	83.4	-23.1	97.5	-0.3	99.4	1.2	90.9	-3.9
vap	67.7	-5.6	77.0	-5.4	76.0	7.4	69.4	1.2	77.6	2.1
var	96.3	2.1	97.3	2.5	97.1	3.3	96.7	5.3	97.2	3.7
vec	70.7	24.4	72.5	-4.3	71.1	15.0	62.8	35.5	80.0	26.3
ven	95.8	5.9	95.7	2.3	96.7	-0.5	95.6	1.4	98.2	-0.1
vep	74.2	-28.7	66.8	-48.3	95.7	-3.8	90.2	15.9	98.2	1.7
vid	75.3	10.8	94.3	-1.1	91.7	0.5	81.5	2.1	91.1	8.6
vie	94.0	-6.9	92.8	-12.1	98.4	-1.0	95.0	-6.4	99.5	-0.2
vif	92.6	3.1	95.6	2.7	91.4	-6.2	91.4	-4.3	92.8	-3.8
viv	95.4	4.4	98.6	-0.2	96.1	3.7	94.8	7.2	96.8	4.4
vls	56.2	-10.6	72.4	-5.9	53.5	-12.8	33.2	60.3	49.6	25.8
vmj	99.1	1.1	99.6	-0.7	98.9	-1.4	99.1	-0.9	98.9	-1.8
vmk	72.3	-8.1	91.8	2.0	88.7	7.6	80.9	12.9	89.2	10.4
vmw	77.1	-13.1	90.5	-16.4	92.4	-3.0	90.2	1.8	94.5	-5.2
vmy	99.0	0.2	99.1	-1.1	98.7	-0.9	98.1	1.2	98.8	-1.4
vol	91.2	1.3	91.2	-7.9	94.5	-0.6	94.6	4.1	97.2	1.3
vro	71.5	22.4	76.8	10.6	56.2	47.0	40.4	67.1	8.1	83.4
vrs	91.4	7.8	96.9	2.5	91.5	-4.7	92.5	-3.0	93.6	-2.8
vun	87.3	7.8	94.8	5.0	89.7	-4.4	87.0	-7.3	86.7	-12.5
vut	96.2	2.6	97.9	2.6	97.7	0.4	95.3	-3.5	96.3	-2.1
waj	98.7	-0.6	99.5	-0.8	99.0	-0.6	98.8	0.2	98.6	-1.3
wal	59.2	-47.2	87.0	-20.8	92.8	2.4	87.4	10.1	92.9	5.8
wap	98.4	-0.1	99.5	-0.0	98.5	1.4	97.9	1.1	98.3	0.6
war	76.3	-2.6	84.4	-3.2	83.4	-5.7	81.1	-0.2	85.4	-9.1
wat	97.2	-3.9	98.6	-2.6	98.0	-2.7	97.4	-3.1	98.4	-2.0
wau	82.4	26.2	75.2	7.3	82.9	27.9	82.9	28.5	83.3	27.9
waw	98.8	0.1	99.3	-1.3	97.5	2.8	96.1	7.1	96.8	5.8
way	95.1	-9.2	96.7	-6.4	99.8	-0.0	99.7	0.1	99.8	0.3
wba	97.1	-5.7	98.9	-2.1	99.7	-0.5	99.8	-0.2	99.8	-0.3
wbi	90.0	2.3	97.4	-5.1	96.5	-0.9	94.5	5.0	96.8	-1.4
wbm	57.5	9.1	39.0	71.6	57.8	0.1	57.8	1.2	56.5	-5.8
wbp	97.7	0.2	98.3	-1.2	98.0	-1.1	97.8	2.8	99.3	1.0
wed	94.9	-0.4	95.9	-4.1	96.4	-0.4	94.4	-0.9	95.8	-5.5
weh	97.1	0.9	98.8	2.1	97.1	0.5	95.0	-1.8	97.5	0.3
wer	97.7	-2.1	98.8	-2.4	98.7	0.5	98.4	2.5	97.2	-2.4
whk	96.5	5.4	95.6	3.6	96.3	1.4	94.9	-1.1	96.9	1.1
wim	97.7	-1.5	98.7	-0.8	98.9	0.6	99.0	1.3	99.0	1.2
win	97.7	-4.5	97.0	-5.7	99.6	-0.8	99.9	0.1	99.7	0.2
wiu	96.9	-4.5	99.3	-0.9	98.9	-0.3	99.0	0.2	99.4	0.1
wiv	95.0	-7.1	98.7	-1.7	99.0	-0.3	97.4	-2.5	98.5	-1.9
wln	94.5	3.3	97.2	-0.5	94.2	0.8	90.6	11.4	96.0	3.9
wls	92.7	3.0	97.4	2.5	95.4	4.6	93.2	1.9	96.3	1.6
wlw	82.0	9.8	90.1	17.5	84.0	2.9	81.5	6.3	84.8	5.2
wlx	96.3	-7.0	99.7	-0.6	99.6	0.1	99.4	1.0	99.8	0.2
wmw	77.0	-2.9	95.0	-2.1	90.9	0.0	78.8	-9.3	90.5	-4.1
wnc	96.5	-1.3	97.6	-4.3	96.8	-1.2	96.5	-0.1	97.3	-2.3
wno	57.2	-4.2	71.3	-4.2	60.1	1.9	60.2	-8.1	62.4	2.4
wnu	96.5	-6.0	99.1	-1.2	99.6	0.2	99.5	0.7	99.6	0.5
wof	96.7	2.7	98.0	2.6	87.3	-18.6	91.9	-9.0	93.1	-9.4
wol	67.9	22.4	80.0	29.3	67.8	16.3	66.2	29.1	74.6	29.5
wos	98.3	-2.7	98.6	-2.3	99.7	-0.1	99.5	0.8	99.5	0.2
wpc	98.2	3.5	98.6	0.6	96.2	5.8	97.4	4.6	97.8	3.8
wrk	98.5	-2.5	99.1	-1.8	99.7	-0.1	99.3	-0.4	99.8	-0.2
wrs	93.8	-10.6	96.8	-6.0	98.2	-2.7	98.7	-0.8	98.5	-2.3
wsg	90.7	-1.5	61.5	43.7	97.9	-1.8	98.0	-0.8	98.9	-1.2
wsk	87.0	-15.6	98.8	-1.7	95.7	-0.0	93.1	6.1	97.0	2.0
wtk	98.1	1.4	96.8	-1.8	97.6	-1.2	98.0	2.0	98.3	1.5
wul	87.7	18.1	90.0	16.1	84.9	7.1	83.5	5.4	85.0	2.7
wuu	57.0	-40.8	52.4	-17.4	58.6	-40.0	79.6	-1.9	88.2	0.0
wuv	96.6	1.1	98.6	0.5	96.9	2.9	96.1	2.3	97.2	3.1
wwa	97.5	-4.6	98.0	-3.8	99.1	-1.6	99.4	-0.2	99.4	0.0
xal	87.1	1.8	95.6	-1.0	95.7	2.0	90.7	0.2	97.1	1.0
xan	91.9	6.1	72.3	35.8	96.7	-1.9	96.8	1.2	97.9	0.6
xav	99.6	-0.7	99.7	-0.4	99.9	-0.1	99.9	0.0	99.9	0.1
xbr	78.3	19.7	91.3	3.4	87.4	6.6	87.2	6.6	90.1	10.0
xcl	83.0	21.8	25.9	83.5	92.6	-0.2	93.2	0.9	97.0	-0.7
xdy	85.7	14.6	92.9	6.6	91.0	-3.1	86.4	-4.5	93.5	-2.8
xed	97.9	1.3	99.3	-0.8	98.2	0.8	97.2	1.6	98.6	0.1
xer	99.6	-0.9	98.7	-2.6	99.9	-0.1	99.9	-0.1	100.0	-0.0
xho	49.7	0.1	81.8	19.1	76.2	-0.4	62.3	3.6	83.9	8.0
xis	92.7	1.0	98.1	1.5	98.0	-2.5	98.0	0.5	98.7	-0.8
xkl	95.3	-7.7	98.5	-2.7	97.6	-2.9	97.3	-4.3	97.9	-4.1
xla	97.9	-2.8	98.8	-1.3	98.9	0.8	98.9	1.4	99.4	0.8
xmf	87.5	6.8	85.6	6.2	94.5	5.0	90.6	15.7	98.2	1.8
xmm	56.3	19.3	83.3	17.3	78.3	8.6	65.8	-4.9	80.3	6.9
xnj	83.6	-9.5	97.3	-3.0	94.6	-3.9	87.9	-5.3	95.3	-2.7
xnn	88.2	-6.2	95.5	-7.4	95.4	-4.0	93.1	-6.9	95.8	-4.0
xno	82.4	-3.4	88.6	-8.6	90.2	-0.6	79.0	-2.8	84.6	-16.4
xnr	62.8	4.5	91.7	12.9	92.5	-0.1	87.3	3.8	92.9	-1.8
xnz	94.2	-4.2	97.9	1.3	98.7	-0.8	98.0	-1.1	98.6	-1.1
xog	71.5	23.6	88.2	13.0	85.9	11.2	77.6	9.7	86.6	12.0
xon	96.7	-4.6	99.5	-0.8	98.1	-2.7	97.1	-3.9	97.3	-4.1
xpe	96.6	0.6	98.6	1.4	97.5	-1.3	97.3	1.6	98.0	2.1
xri	99.7	-0.7	99.6	-0.8	99.9	-0.1	99.7	-0.6	100.0	-0.1
xsb	90.1	-0.5	96.1	-5.8	96.1	-4.8	94.9	-5.6	98.3	-1.9
xsi	98.9	-0.8	98.6	-2.4	99.7	0.3	99.5	0.9	99.6	0.4
xsm	97.0	2.5	98.6	2.4	96.5	1.0	96.6	0.8	97.2	-0.9
xsu	95.9	7.1	96.0	4.3	95.3	5.1	95.2	5.0	95.7	4.6
xta	84.5	-20.8	95.2	-5.9	96.1	-3.6	94.7	-4.2	98.4	-1.2
xtd	97.3	3.5	99.3	0.4	97.5	-0.4	97.6	1.5	98.4	0.7
xte	97.3	-3.3	98.4	-1.2	98.8	-0.4	98.0	-1.2	99.4	0.2
xtm	95.6	-5.5	99.5	0.1	98.9	-0.3	97.7	-1.2	99.1	-0.1
xtn	94.2	2.9	97.8	2.1	96.6	1.4	92.3	8.6	97.6	1.7
xuo	98.2	1.7	98.6	1.7	99.2	0.4	99.3	0.6	99.3	-0.4
yaa	99.1	-0.4	98.8	-1.5	99.2	0.5	99.5	0.9	99.4	0.7
yac	94.0	2.4	97.2	-1.8	97.2	-2.0	95.4	-1.9	97.4	-1.8
yad	99.2	-0.2	99.5	-1.1	99.3	0.7	99.4	1.0	99.5	0.9
yae	99.5	1.0	98.8	-0.6	99.5	0.9	99.5	0.9	99.2	0.3
yal	73.9	-22.3	87.8	-18.8	96.4	1.9	93.5	4.5	96.6	3.2

Table 9: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.

Lang	Textcat		NB		fastText		LSTM		GLOT500	
	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec	F1	Prec-Rec
yam	98.5	0.2	99.9	-0.1	98.6	-0.2	98.4	1.1	99.1	-0.3
yan	92.7	-0.1	97.2	3.4	94.6	-0.1	92.8	-3.5	97.0	1.3
yao	87.5	10.9	92.1	6.3	90.1	5.3	85.8	2.3	90.9	4.9
yap	89.7	5.0	94.3	8.4	92.9	11.6	92.9	11.0	94.3	9.7
yaq	97.8	-2.5	99.6	-0.5	98.8	0.7	98.5	0.9	99.1	-0.6
yas	92.0	3.4	95.8	0.8	95.3	0.9	94.7	0.4	93.5	-2.6
yat	96.2	-1.6	92.5	-13.4	96.4	3.0	96.6	2.0	98.0	0.7
yaz	99.1	-1.7	99.5	-1.0	99.8	-0.3	99.9	0.1	99.9	-0.2
yba	96.6	5.1	96.4	3.5	96.8	6.0	97.0	5.5	97.7	4.3
ybb	96.3	6.5	97.0	4.5	96.4	-2.3	97.4	-1.7	96.7	-3.4
ybh	46.2	46.5	33.4	-44.4	44.7	-67.7	44.0	-63.7	44.6	-68.4
yby	86.3	-3.0	91.8	7.9	95.0	-5.8	97.0	-2.4	98.5	-0.2
ycl	99.2	-1.5	99.5	-1.1	99.6	-0.6	99.9	-0.2	99.8	-0.4
ycn	97.8	-3.3	99.6	-0.5	99.1	-0.6	99.2	0.1	99.4	-0.3
ydd	99.2	0.5	98.1	-1.8	99.9	0.0	99.6	0.6	99.9	-0.1
yea	95.5	-2.6	98.7	-1.5	99.4	0.6	98.8	-0.6	99.4	-0.3
yim	93.1	7.1	94.9	5.9	92.4	3.9	89.5	0.8	94.2	7.0
yka	98.7	0.2	98.7	-1.7	98.5	-0.6	98.8	1.5	98.9	2.1
yle	99.2	1.4	99.7	-0.4	98.5	1.6	99.1	1.0	98.8	1.0
yli	88.6	1.1	94.0	2.6	93.7	4.4	88.2	0.1	91.7	-1.9
yml	98.4	0.0	98.7	0.1	97.6	-1.5	96.6	0.5	98.3	-0.5
yom	77.0	16.3	89.6	8.5	90.9	3.2	86.6	-0.1	94.8	0.2
yon	97.5	-4.7	98.8	-2.1	99.5	-0.5	99.5	-0.5	99.5	-0.7
yor	90.4	11.2	96.9	3.7	94.3	-1.5	95.1	2.3	97.1	0.4
yrb	95.7	-2.3	98.5	-0.7	98.0	-0.6	96.1	-0.5	98.8	0.4
yre	95.8	-4.0	99.1	-1.5	97.6	1.4	97.2	2.8	98.9	1.2
yrl	96.8	2.9	98.8	0.8	96.9	-1.0	95.7	1.3	97.0	-1.6
yua	94.0	6.4	94.6	1.7	91.2	-3.9	90.5	-4.8	94.3	-1.8
yue	70.7	27.2	54.7	59.2	72.1	4.7	94.1	5.0	97.3	-0.2
yuj	95.5	-5.1	99.0	-1.5	98.4	1.2	98.0	1.9	98.6	2.0
yut	99.2	-0.3	99.4	-0.8	99.6	-0.2	99.1	-0.1	99.4	-0.9
yuw	98.6	-1.7	98.2	-3.4	98.7	-1.6	98.9	-1.0	99.5	-0.1
yuz	99.1	-1.2	96.3	-7.1	99.6	-0.4	99.7	0.1	99.7	-0.3
yva	96.0	1.0	98.2	1.4	97.0	2.0	95.7	1.3	97.2	1.9
zaa	98.9	-0.6	99.2	-1.1	98.9	-0.8	98.8	-1.1	99.5	-0.6
zab	96.8	0.2	98.7	0.5	97.6	-2.2	97.6	-1.9	98.6	-1.2
zac	98.3	0.9	99.0	-0.2	98.8	-0.2	98.5	-0.9	99.6	0.3
zad	95.9	-5.0	96.9	-5.0	96.2	-3.7	95.6	-1.5	98.3	-0.6
zae	96.8	-0.2	98.7	-0.4	97.4	0.0	96.0	0.0	97.8	1.5
zai	94.4	-3.2	98.3	0.1	97.2	0.6	96.1	-1.4	98.1	-0.4
zaj	40.8	-6.2	79.5	-7.2	75.3	-6.4	58.8	-1.7	75.6	-9.8
zam	95.2	-5.7	98.9	-2.0	98.2	-0.4	97.9	0.1	99.1	0.9
zao	91.7	-3.2	96.2	1.2	94.7	-0.3	93.3	1.9	95.9	1.2
zaq	94.6	3.1	97.2	2.7	95.1	0.5	94.9	0.7	96.9	-0.2
zar	95.6	7.3	88.3	18.0	97.6	1.3	98.8	0.9	99.1	-0.2
zas	97.8	0.7	98.9	0.6	97.9	-0.3	97.2	-1.1	98.7	-0.1
zat	93.4	10.7	95.1	7.2	95.8	5.8	96.0	3.5	98.0	0.6
zav	99.3	0.5	98.6	-2.3	99.0	0.2	98.9	1.2	99.6	0.1
zaw	98.4	0.1	99.0	-1.0	98.0	-1.1	98.2	0.6	99.1	-0.1
zca	96.9	3.7	98.8	0.7	97.0	3.5	95.9	0.4	98.1	2.1
zea	19.9	2.9	58.3	-2.9	32.4	17.9	4.9	50.2	50.6	34.0
zga	90.4	11.9	95.8	1.7	92.6	7.6	90.0	15.2	93.7	8.4
zgh	79.9	-33.5	79.4	-32.0	83.0	-13.1	76.1	-4.3	81.8	-11.5
zhn	89.6	13.1	87.2	21.3	94.2	2.2	90.5	-3.6	94.8	-2.8
zia	98.2	-1.9	99.6	-0.3	98.8	0.2	98.2	2.5	98.9	1.3
zim	97.6	-3.9	99.8	-0.4	99.5	-0.3	99.3	-0.3	99.8	-0.3
ziw	72.9	-17.0	93.1	-4.8	88.6	-4.2	77.3	-10.9	86.6	-9.9
zlm	47.4	0.6	68.0	-4.9	61.9	-3.6	50.2	-7.6	63.8	-16.3
zmb	87.9	3.6	94.9	5.7	80.9	-21.8	85.2	-13.1	86.2	-15.9
zmz	99.8	-0.4	99.6	-0.7	99.8	-0.4	99.9	0.2	99.8	0.1
zne	91.1	3.5	96.3	5.2	96.1	2.1	95.2	0.4	97.6	0.5
zoc	98.8	2.1	98.6	1.1	97.1	-2.7	95.6	-6.1	96.4	-5.5
zom	73.8	15.7	85.3	5.4	85.1	11.8	75.0	10.1	81.8	12.8
zos	99.6	0.7	99.5	-0.4	98.9	-1.4	99.4	-0.1	99.6	-0.0
zpc	96.7	-2.8	98.2	-3.3	98.9	-0.1	98.8	0.9	99.5	0.0
zpg	96.5	6.2	98.6	1.3	96.9	2.0	96.2	2.2	98.1	2.5
zpi	96.4	-1.8	98.1	-0.2	96.7	-1.4	95.3	-2.9	97.5	-1.5
zpl	97.9	2.5	98.8	0.8	98.2	0.8	97.7	2.2	98.5	0.0
zpm	98.0	0.9	98.6	0.2	97.0	1.2	96.5	0.9	98.1	1.2
zpo	93.9	3.8	96.4	2.5	97.4	1.4	97.2	2.0	98.4	1.2
zpq	88.2	-20.2	94.3	-10.1	91.8	-13.9	84.4	-25.6	96.5	-5.4
zpt	97.8	-0.4	99.3	-0.4	98.4	-0.7	98.0	1.5	98.9	0.3
zpu	98.7	-0.4	99.5	-0.5	98.9	1.5	98.6	0.1	99.1	0.6
zpv	99.3	0.0	99.4	-0.7	99.4	-0.2	99.2	1.0	99.3	0.8
zpz	98.3	-1.2	99.0	-1.5	99.0	0.7	98.2	-0.3	99.3	0.3
zrs	82.5	-0.5	89.7	9.6	83.6	-0.4	85.1	11.2	87.2	8.0
zsm	42.5	7.2	56.3	11.8	53.6	13.6	45.2	16.8	47.3	22.6
zsr	96.3	-1.4	91.6	-13.3	97.6	-0.2	97.6	0.7	98.8	-0.3
ztg	98.2	1.0	98.9	0.0	98.7	0.2	98.6	0.7	99.3	0.4
ztq	97.6	-0.8	99.2	-0.2	98.4	0.4	98.5	0.4	99.1	-0.3
zty	79.2	32.4	88.7	18.3	84.8	22.2	76.5	34.9	94.9	7.3
zul	46.8	-6.7	75.4	-2.6	76.3	-2.8	57.7	13.9	80.0	-2.5
zyb	94.8	3.1	91.2	-7.6	94.7	5.6	93.3	3.0	96.1	4.9
zyp	97.8	0.6	98.4	-1.4	98.0	1.7	95.8	1.2	98.4	1.7

Table 10: Results per language of the model with all 2,034 languages in our benchmarks. We report F1 score, and precision-recall.