cantnlp@LT-EDI-2024: Automatic Detection of Anti-LGBTQ+ Hate Speech in Under-resourced Languages

Sidney G.-J. Wong[1,2] and Matthew Durward[1]

[1]University of Canterbury, New Zealand
[2]Geospatial Research Institute, New Zealand
{sidney.wong,matthew.durward}@pg.canterbury.ac.nz

Abstract

This paper describes our homophobia/transphobia in social media comments detection system developed as part of the shared task at LT-EDI-2024. We took a transformer-based approach to develop our multiclass classification model for ten language conditions (English, Spanish, Gujarati, Hindi, Kannada, Malayalam, Marathi, Tamil, Tulu, and Telugu). We introduced synthetic and organic instances of script-switched language data during domain adaptation to mirror the linguistic realities of social media language as seen in the labelled training data. Our system ranked second for Gujarati and Telugu with varying levels of performance for other language conditions. The results suggest incorporating elements of paralinguistic behaviour such as script-switching may improve the performance of language detection systems especially in the cases of under-resourced languages conditions.

1 Introduction

The purpose of this shared task was to develop a multiclass classification system to predict instances of homophobia/transphobia in social media comments across different language conditions (Kumaresan et al., 2024). The ten language conditions were: English (ENG), Spanish (ESP), Gujarati (GUJ), Hindi (HIN), Kannada (KAN), Malayalam (MAL), Marathi (MAR), Tamil (TAM), Tulu (TCY), and Telugu (TEL).

The main contribution of this paper is that we extend on the work using spatio-temporally retrained transformer-based language models in Wong et al. (2023). We have expanded on the synthetic script-switching approach by incorporating real-world (or organic) samples of script-switching during domain adaptation in the development of our multiclass classification model using pretrained language models.

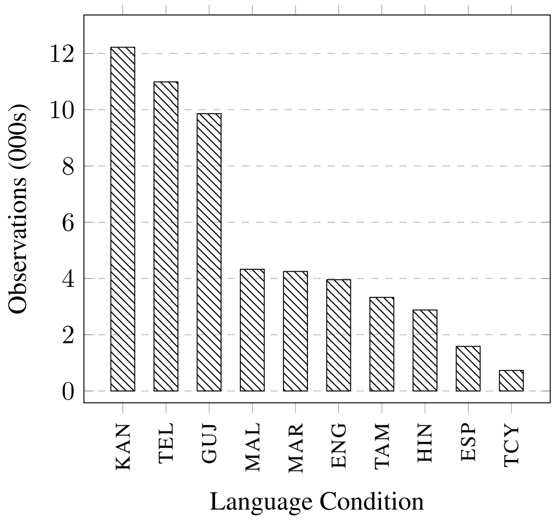

Figure 1: Barplot of labelled training data. The combined total number of observations (in thousands) by language condition ordered from the most (KAN) to the least (TCY) number of observations.

1.1 Problem Description

The organisers of the shared task provided labelled training data for each of the ten language conditions. Five of the language conditions belong to the Indo-European language family (ENG, ESP, GUJ, HIN, and MAR) and the remaining five language conditions belong to the Dravidian language family (KAN, MAL, TAM, TCY, and TEL).

The labelled training data comes from different sources (ENG and TAM in Chakravarthi et al., 2021; HIN and MAL in Kumaresan et al., 2023; and ESP in García-Díaz et al., 2020). The training data is made up of comments from users reacting to LGBTQ+ related content on YouTube. The labelled training data for GUJ, KAN, MAR, TCY, and TEL were introduced for the current shared task.

The total number of social media comments for each language condition (combining the train and development sets) are shown in Figure 1. KAN has the most observations, followed by TEL and GUJ.

Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity, Inclusion, pages 177–183

	NONE	HOMO	TRANS
ENG	0.94	0.06	0.00
ESP	0.57	0.22	0.22
GUJ	0.47	0.28	0.25
HIN	0.95	0.02	0.04
KAN	0.44	0.27	0.28
MAL	0.79	0.16	0.06
MAR	0.73	0.16	0.11
TAM	0.77	0.17	0.06
TCY	0.74	0.26	-
TEL	0.39	0.32	0.30

Table 1: Class distribution by language condition. Note that TCY has a binary class distribution.

TCY has the least number of observations. The remaining language conditions each have between 1,000 to 5,000 observations.

The social media comments were manually annotated and broadly labelled using on a three-class classification system (Chakravarthi et al., 2021). There were only two classes for TCY which we have labelled NONE and HOMO for consistency with other language conditions. The classes are:

- *Homophobic Content* (HOMO): any comments which were deemed gender-based and involved pejorative or defamatory language directed towards non-heterosexual people.
- *Transphobic Content* (TRANS): any derogatory or offensive language directed towards transgender and gender diverse people.
- *Non-anti-LGBTQ+ Content* (NONE): counter speech or hope speech as well as comments which does not contain any homophobic or transphobic content.

The class distribution for each language condition is shown in Table 1. We observe significant class imbalance between language conditions especially in ENG and HIN where the HOMO and TRANS classes make up less than a tenth of the labelled training data. Of the 3,726 observations in the ENG language condition, there are only 221 tokens of HOMO and nine tokens of TRANS.

Outside the labelled training data and published material, the organisers did not provide additional corpus or demographic information of the labelled training data as part of the shared task. Therefore, the classification system needs to account for the differences in data availability as well as class imbalance for each language condition.

1.2 Related Work

The current shared task is the third shared task on homophobia and transphobia detection in social media comments. The first shared task involved only three language conditions: TAM, ENG, and a separate TAM-ENG code-mixed condition (Chakravarthi et al., 2022).

The classification system with the best performance for ENG had a weighted Macro F_1 score of 0.92 was developed by team ABLIMET (Maimaiti-tuoheti et al., 2022) and for TAM was 0.94 developed by team ARGUABLY. The best performing classification system for the TAM-ENG code-mixed condition was also developed by team ARGUABLY with a weighted Macro F_1 score of 0.89. The code-mixed condition had the lowest performance across the three conditions.

Participants took different approaches involving statistical language models and machine learning. The best performing system used XLM-ROBERTA pretrained language models (Conneau et al., 2020). This BERT-based transformer language approach structures the relationship between words with language embeddings (Devlin et al., 2019). These language embeddings account for structures across multilingual conditions.

The second shared task expanded to five language conditions (ENG, ESP, HIN, MAL, and TAM) which was broken down by a three-class classification system similar to the current shared task (Chakravarthi et al., 2023). Three of the language conditions (ENG, MAL, and TAM) were further classified into a seven-class classification system.

The weighted Macro F_1 score for the best performing three-class classification systems was 0.97 for ENG and 0.98 for HIN developed by TEAMPLUSONE using BERT-based transformer models. A weight-space ensembling technique presented itself as the best solution for ESP, MAL, and TAM language conditions (Ninalga, 2023).

The best performing systems for the seven-class classification condition were all developed using transformer language models. The weighted Macro F_1 score ENG was 0.82 developed by team TEAMPLUSONE, for MAL was 0.88 developed by team CANTNLP (Wong et al., 2023), and for TAM was 0.87 developed by team DEEPBLUEAI.

This suggests BERT-based models, such as XLM-ROBERTA for zero-shot learning, are particularly effective in carrying out multiclass classification tasks outlined in the current shared task. More

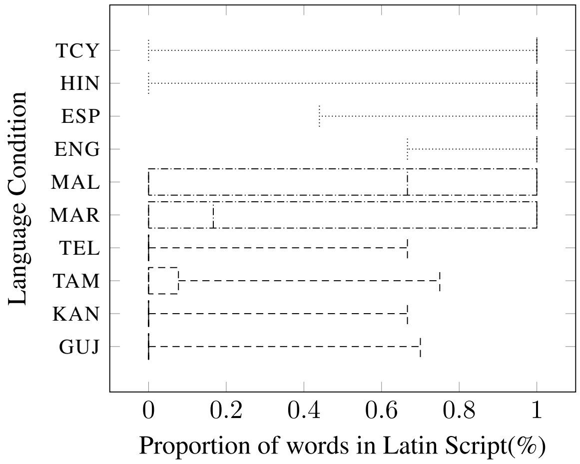

Figure 2: Boxplot of labelled training data. Language condition by the proportion of observations with at least one word written in Latin script ordered from the lowest (TCY) to the highest (GUJ) proportion of observations.

importantly, these systems are simple to implement and allow for domain adaptation (Liu et al., 2019)

Wong et al. (2023) introduced synthetically script-switched instances of social media data during domain adaptation to account for the high frequency of script-switching in the labelled data for HIN, MAL, and TAM. The introduction of script-switched language data improved the performance of the homophobia/transphobia detection model in HIN, but not MAL or TAM.

The results from Wong et al. (2023) suggest that there is potential for incorporating paralinguistic behaviour such as script-switching in the development of multiclass detection language systems. Therefore, this paper explores this further by incorporating different forms of script-switching.

2 Methodology

In this section, we provide an overview of our system development methodology. We took a transformer-based language model approach to develop our system. We used XLM-ROBERTA as the base PLM for our system (Conneau et al., 2020). The embeddings in XLM-ROBERTA were trained on two terabytes of web-crawled data for over 100 language including nine of the ten language conditions of interest (with the exclusion of TCY).

A significant advantage of transformer-based PLMs is the ability for domain adaptation as discussed in Section 2.1. This means we can retrain the default language embedding models with additional language data without the need to train resource-intensive PLMs from scratch.

We tested different forms of script-switching in order to understand the impacts of script-switching on our classification system. We then used the PLMs developed in Section 2.1 to fine-tune our multiclass classification model as discussed in Section 2.2. Based on the weighted Macro F_1 for each language condition, we submitted the results from the best performing multiclass classification system to the organisers.

2.1 Domain Adaptation

The first stage in developing our system involved domain adaptation (also known as retraining). Liu et al. (2019) noted that domain adaptation can improve the performance of transformer-based language models in downstream tasks. We can do this by introducing domain (or register) specific text samples to produce customised retrained PLMs (or retrained language models). This means we can introduce language data from under-resourced languages such as TCY as well as additional linguistic information such as script-switching - a common phenomenon in social media language.

As noted in Wong et al. (2023), we observed varying levels of script-switching in the labelled training data. Therefore, we first needed to identify the level of script-switching between language conditions. For each observation, we calculated the proportion of words written in Latin script using the alphabet-detector[1] Python package.

The proportion of script-switching between language conditions is shown in Figure 2 where 0 suggests low usage of Latin-based characters (in the case of GUJ, KAN, TAM, TEL) while 1 suggests high usage as expected for ENG and ESP. Figure 2 confirms that there is sufficient need to account for the varying-degrees of script-switching between language conditions.

We retrained XLM-ROBERTA with two forms of script-switching: synthetic and organic script-switching. These are our candidate models. We describe how we produced the language data for domain adaptation in Section 2.1.1 and Section 2.1.2. We produced the candidate language models by retraining the language embeddings using the simpletransformers[2] Python library. We did this over four iterations and we evaluated the training for every 500 steps with AdamW optimisation (Loshchilov and Hutter, 2019). Model performance was based on evaluation loss.

[1]https://pypi.org/project/alphabet-detector/
[2]https://simpletransformers.ai/

	BASELINE		SYNTHETIC		ORGANIC	
	mono	*multi*	*mono*	*multi*	*mono*	*multi*
ENG	0.32	**0.35**	0.32	0.32	0.32	0.32
ESP	0.80	0.82	**0.87**	0.84	0.76	0.82
GUJ	0.94	0.94	**0.95**	**0.95**	**0.95**	**0.95**
HIN	**0.32**	**0.32**	**0.32**	**0.32**	**0.32**	**0.32**
KAN	0.92	0.93	**0.94**	**0.94**	**0.94**	**0.94**
MAL	0.51	0.53	0.73	0.58	**0.78**	0.61
MAR	0.44	0.45	0.44	0.41	0.42	**0.46**
TAM	0.48	0.42	0.54	0.49	0.48	0.56
TCY	**0.72**	0.43	0.43	0.43	0.43	0.43
TEL	**0.98**	0.97	**0.98**	0.97	0.97	**0.98**

Table 2: Model performance of candidate classification models by Macro F_1 using our test set split from combining the train and validation sets provided to us by the organisers for each language condition. The three candidate languages models are: BASELINE, SYNTHETIC, and ORGANIC. We also compared the performance of language-specific (*mono*) and multilingual (*multi*) multiclass classification models. The best performing system is highlighted in **bold**.

2.1.1 Synthetic Script-Switching

We took a similar approach as Wong et al. (2023) to produce synthetic samples of script-switched language data for domain adaptation. We define synthetic as machine-generated texts. Due to the limited availability of observations for some language conditions, our main source of human-generated texts come from the Leipzig Corpus Collection (Goldhahn et al., 2012). Each corpus contained 10,000 Wikipedia abstracts produced in 2016 with the exception of TCY which was produced in 2018.

We then randomly sampled half of the abstracts from each language condition (excluding the Latin-based ENG and ESP) and used the ai4bharat[3] Python library to transliterate the relevant Brahmic orthographies into Latin script. Once we produced a subset of synthetically script-switched Wikipedia abstracts, we combined the original abstracts with the synthetically script-switched abstracts. Finally, we combined the labelled training data to create train and evaluation sets. The inclusion of the labelled training data is to ensure register-specific domain adaptation.

2.1.2 Organic Script-Switching

The second form of script-switched language data for domain adaptation involve organic samples of script-switched language data. We define organic as human-generated texts. This proved to be a challenge as we were unable to identify sources of script-switched social media language data for some of the under-resourced language conditions.

We used the pre-existing labelled training data to produce language profiles to develop a language identification model with the langdetect[4] These language profiles were used to detect organic instances of script-switched social media data from the Global Corpus of Language Use (CGLU; Dunn, 2020). This produced a train set with 230,000 observations and an evaluation set with 12,000 observations which we could use for domain adaptation.

2.2 Classification Model

As discussed in Section 2.1, we developed our multiclass classification models using the candidate language models during the domain adaptation phase. The three candidate languages models are: the baseline XLM-ROBERTA language model (BASELINE), XLM-ROBERTA retrained with synthetic samples of script-switched language data (SYNTHETIC), and XLM-ROBERTA retrained with organic samples of script-switched language data (ORGANIC).

We resampled the available data to create our own train (80%), validation (10%), and test (10%) sets to avoid over-fitting on the validation set during model evaluation. We trained language specific classification models (*mono*) and an ensemble multilingual classification model (*multi*) by combining the labelled training data. We trained the multiclass classification model using the simpletransformers Python package for four iterations and we evaluated the training for every 500 steps with AdamW optimisation (Loshchilov and Hutter, 2019). The model performance was based on evaluation loss.

[3] https://pypi.org/project/ai4bharat-transliteration/

[4] https://pypi.org/project/langdetect/

	CANTNLP	Best Performance
ENG	0.323	*0.496*
ESP	0.496	*0.582*
GUJ	0.962	*0.968*
HIN	0.326	*0.458*
KAN	0.943	*0.948*
MAL	0.775	*0.942*
MAR	0.433	*0.626*
TAM	0.555	*0.880*
TCY	0.452	*0.707*
TEL	0.965	*0.971*

Table 3: The average Macro F_1 score of our classification system, and the average Macro F_1 score of the overall best performing classification system.

The model performance for each of the candidate models are shown in Table 2. We have indicated the best performing model based on average Macro F_1 score (highlighted in **bold**). In some language conditions, there were multiple best performing models. Not included in Table 2 are the combined average Macro F_1 score for the multilingual models: the average Macro F_1 for the BASELINE model was 0.89; both SYNTHETIC and ORGANIC models had an average Macro F_1 of 0.90.

3 Results

Based on the average Macro F_1 score of the candidate models as shown in Table 2, we nominated the language-specific synthetic classification model as the best performing classification system. We applied this classification system and submitted the results to the organisers. The results of our submitted homophobia/transphobia detection system are shown in Table 3. The best performing language condition was TEL with an average Macro F_1 of 0.97 and the worst performing language condition was ENG with an average Macro F_1 of 0.32.

Our final rank for each language conditions are as follows: for ENG we came tenth equal out of ten teams; for ESP we came fourth out of five teams; for GUJ we came second out of six teams; for HIN we came fourth equal out of seven teams; for KAN we came fourth out of eight teams; for MAL we came seventh out of nine teams; for MAR we came fifth out of six teams; for TAM we came fifth out of of eight teams; for TCY we came third equal out of four teams; and finally for TEL we came second out of nine teams.[5]

[5]Note the final rankings differ from the published results for ESP, TAM, and TEL as they were not included in the final rank list due to human error from the organising committee during submission.

4 Discussion

The use of synthetic and organic script-switched language data during domain adaptation increased the performance for all language conditions from the BASELINE model with the exception of ENG, HIN, and TCY. We expected the ENG and ESP language conditions to perform poorly with our proposed methodology as there were very few instances of script-switching, but the poor performance of TCY was unexpected.

We hypothesise the poor performance in TCY was due to the limited number of observation (as shown in Figure 1) the higher than expected usage of Latin-based script for TCY in the labelled training data (as shown in Figure 2). This will require robust statistical analysis beyond the scope of the current paper.

We also posit the poor performance of ENG and HIN was a result of the class imbalance between instances of homophobic, transphobic and the non-anti-LGBTQ+ content as demonstrated in Table 1. The performance of our ENG and HIN language-specific detection models are in line with other participating teams.

In contrast to the method proposed in Wong et al. (2023), we did not include any methods to counter the class imbalance in the training data nor did we include random noise injection to expand the minority classes. It was shown that random over sampling of minority classes did not significantly improve the performance of the detection models.

5 Conclusion

The main contribution of the current paper is the proposal to use synthetic and organic script-switching examples of during domain adaptation to improve the down-stream performance for under-resourced languages. We demonstrated that our methodology improved the model performance for GUJ, KAN, MAL, MAR, and TAM even though the improvement was only marginal. Even though our homophobia/transphobia detection system did not rank first for any of the ten language conditions, we were pleased with the performance of our detection system which supports the inclusion of paralinguistic information.

Ethics Statement

The purpose of the current shared task is to develop a homophobic/transphobic language detection system in social media texts particularly for under-resourced Indo-Aryan and Dravidian languages within the fields of computational linguistics and natural language processing.

We recognise the importance of community-lead research in particular by members of under-represented and minoritised communities. The lead author acknowledges his positionality as a member of the LGBTQ+ community. The lead author is familiar with anti-LGBTQ+ discourse both in online and offline spaces and the harmful effects of hate speech and offensive language on members of the LGBTQ+ communities (Wong, 2023b).

In terms of the authors' linguistic membership, the authors share proficiency in ENG and ESP; however, the authors acknowledge their limited experience with GUJ, KAN, MAR, TAM, TCY, and TEL with some exposure to HIN and MAL. We acknowledge the limitations of our analysis in language conditions where we have limited proficiency and we will follow the guidance and expertise of members from the relevant language communities.

We want to thank the organisers of the shared task and the workshop on Language Technology for Equality, Diversity, and Inclusion. We also want to thank the contributors of the training data and those who were involved in the labelling process across the different language conditions.

Limitations

Under the purview of developing a homophobic/transphobic language detection system in social media texts, we want to highlight the limitations of our proposed system and methodology.

Firstly, we acknowledge there are differences in data quality and veracity between the different language conditions. This is based on the differences in the corpus size between the different language conditions (as shown in Figure 1) and the distribution of homophobic and transphobic content.

In light of these data quality issues, we have not accounted for these differences between language conditions. This means we do not entirely understand the downstream impacts on model performance - although it is clear that there is a possible relationship between larger and more balanced language conditions (TEL) performing better than smaller and more imbalanced language conditions (TCY). It is possible these differences could exacerbate biases already observed in transformer-based language models (Bhardwaj et al., 2021).

Beyond the upstream and downstream impacts of bias in transformer-based language models, we also recognised that incorporating external data sets from LCC (Goldhahn et al., 2012) and the CGLU (Dunn, 2020) introduces additional biases not properly addressed in this paper such as geographic bias in social media language data (Wong et al., 2022).

Secondly, there is a need to conduct this form of research under a sociolinguistic or linguistic anthropological framework. There is a risk that training data detecting homophobia, transphobia, hate speech, or offensive may not necessarily reflect the social, political, or linguistic realities of different populations. This is because some of the features extracted from the labelled training data may not reflect real-world knowledge.

These differences are particularly evident when we apply these detection systems across dialect contexts (Wong, 2023a). For this reason, we propose that future work in this area should also consider how these systems perform in real-world context beyond the evaluation of labelled training data. We should work alongside members of LGBTQ+ communities from culturally and linguistically diverse backgrounds to understand the effectiveness and generalisability of our homophobic/transphobic detection systems.

References

Rishabh Bhardwaj, Navonil Majumder, and Soujanya Poria. 2021. Investigating Gender Bias in BERT. *Cognitive Computation*, 13(4):1008–1018.

Bharathi Raja Chakravarthi, Rahul Ponnusamy, Malliga S, Paul Buitelaar, Miguel Ángel García-Cumbreras, Salud María Jiménez-Zafra, Jose Antonio García-Díaz, Rafael Valencia-García, and Nitesh Jindal. 2023. Overview of Second Shared Task on Homophobia and Transphobia Detection in Social Media Comments. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 38–46, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Thenmozhi Durairaj, John McCrae, Paul Buitelaar, Prasanna Kumaresan, and Rahul Ponnusamy. 2022. Overview of The Shared Task on Homophobia and Transphobia Detection in Social Media Comments. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 369–377, Dublin, Ireland. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kayalvizhi Sampath, Durairaj Thenmozhi, Sathiyaraj Thangasamy, Rajendran Nallathambi, and John Phillip McCrae. 2021. Dataset for Identification of Homophobia and Transophobia in Multilingual YouTube Comments. ArXiv:2109.00227 [cs].

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised Cross-lingual Representation Learning at Scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. ArXiv:1810.04805 [cs].

Jonathan Dunn. 2020. Mapping languages: the Corpus of Global Language Use. *Language Resources and Evaluation*, 54(4):999–1018.

José Antonio García-Díaz, Ángela Almela, Gema Alcaraz-Mármol, and Rafael Valencia-García. 2020. UMUCorpusClassifier: Compilation and evaluation of linguistic corpus for Natural Language Processing tasks. *Procesamiento del Lenguaje Natural*, 65(0):139–142.

Dirk Goldhahn, Thomas Eckart, and Uwe Quasthoff. 2012. Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 759–765, Istanbul, Turkey. European Language Resources Association (ELRA).

Prasanna Kumar Kumaresan, Rahul Ponnusamy, Ruba Priyadharshini, Paul Buitelaar, and Bharathi Raja Chakravarthi. 2023. Homophobia and transphobia detection for low-resourced languages in social media comments. *Natural Language Processing Journal*, 5:100041.

Prasanna Kumar Kumaresan, Ruba Priyadharshini, Bharathi Raja Chakravarthi, Paul Buitelaar, Asha Hegde, Hosahalli Lakshmaiah Shashirekha, Saranya Rajiakodi, Miguel Ángel García-Cumbreras, Salud María Jiménez-Zafra, José Antonio García-Díaz, Rafael Valencia-García, Kishore Kumar Ponnusamy, Poorvi Shetty, and Daniel García-Baena. 2024. Overview of Third Shared Task on Homophobia and Transphobia Detection in Social Media Comments. In *Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity and Inclusion*, Malta. European Chapter of the Association for Computational Linguistics.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. ArXiv:1907.11692 [cs].

Ilya Loshchilov and Frank Hutter. 2019. Decoupled Weight Decay Regularization. ArXiv:1711.05101 [cs, math].

Abulimiti Maimaitituoheti, Yong Yang, and Xiaochao Fan. 2022. ABLIMET @LT-EDI-ACL2022: A Roberta based Approach for Homophobia/Transphobia Detection in Social Media. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 155–160, Dublin, Ireland. Association for Computational Linguistics.

Dean Ninalga. 2023. Cordyceps@LT-EDI: Patching Language-Specific Homophobia/Transphobia Classifiers with a Multilingual Understanding. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 185–191, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Sidney Wong, Jonathan Dunn, and Benjamin Adams. 2022. Comparing Measures of Linguistic Diversity Across Social Media Language Data and Census Data at Subnational Geographic Areas. *Proceedings in New Zealand Geospatial Research Conference*.

Sidney Wong, Matthew Durward, Benjamin Adams, and Jonathan Dunn. 2023. cantnlp@LT-EDI-2023: Homophobia/Transphobia Detection in Social Media Comments using Spatio-Temporally Retrained Language Models. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 103–108, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Sidney Gig-Jan Wong. 2023a. Monitoring Hate Speech and Offensive Language on Social Media. In *Fourth Spatial Data Science Symposium*, University of Canterbury.

Sidney Gig-Jan Wong. 2023b. *Queer Asian Identities in Contemporary Aotearoa New Zealand: One Foot Out of the Closet*. Lived Places Publishing.