

HACS-TL: Cross-Script Transfer Learning for Hausa Ajami Hate Speech Detection Using Transformer-Based Architecture

Abdulkadir Shehu Bichi^{1,*}, Jyoti Shekhawat¹, Muquaddar Ali¹
Prashant Sharma¹ Ismail Dauda Abubakar^{1,3}

¹Vivekananda Global University Jaipur, ²Federal University Gusau

Correspondence: abdulkadir.bichi@babaahmeduniversity.edu.ng, 24wtec3csml001@vgu.ac.in

Abstract

Hausa is one of the most widely spoken languages in West Africa and uses two scripts: the Boko Latin and the Ajami Arabic-derived script. Although Ajami is culturally important, it is still very underrepresented in natural language processing. In this paper, HACS-TL, a novel transformer-based architecture for detecting hate speech in Hausa Ajami, is proposed. Our model uses a linguistically motivated Latin-to-Hausa Ajami converter, cross-script attention, and dialect-specific pooling. After conducting extensive 2-fold cross-validation on 2,000 samples, we found that HACS-TL outperformed baseline models (mBERT (73.70%), XLM-RoBERTa (67.20%), and AraBERT (56.96%)) with a macro F1 score of 75.33%. Detailed error analyses and orthographic stress tests demonstrate the robustness of this approach. The proposed Hausa Ajami conversion system achieved 63.57% character-level similarity with 51% correct conversions, thereby establishing a baseline for the computational processing of Ajami Hausa.

Keywords: Cross-script transfer learning, Hate speech detection, and Hausa Ajami

1 Introduction

The current socially mediated global environment has fostered the global spread of hate speech, which has been particularly damaging in regions with limited language moderation tools, such as in many areas of diverse language (Caselli et al., 2021; Fortuna and Nunes, 2018). Despite the unprecedented advancements achieved in the detection of hate speech in major languages, such as English, Arabic, and Hindi (Mozafari et al., 2020; Aluru et al., 2020), the languages of Africa, which are on the lower end of the resource scale, have been grossly ignored (Adelani et al., 2022; Alabi et al., 2022). There are around 77 million speakers of Hausa, a West African language, and most of them pose

several difficulties for NLP tools. Hausa, being a pluricentric language, has a number of different standardization bodies and thus has a great deal of dialectal diversity. Hausa also has two separate orthographies: one based on a Latin script (Boko) and an Arabic-derived script (Ajami) (Newman, 2000; Jagggar, 2001). Although still used in many religious and informal contexts, the Hausa Ajami script has been especially neglected in the field of computational linguistics, despite its cultural relevance (Mumin and Versteegh, 2014). Transformers have been a big leap for cross-lingual and multilingual understanding (Devlin et al., 2019; Conneau et al., 2020) but still largely pay no attention to certain languages and scripts, especially morphologically rich languages and those with non-Latin scripts (Rust et al., 2021; Ponti et al., 2020). Despite demonstrating the value of bottom-up, script-specific pretraining, Arabic-centric models like AraBERT (Antoun et al., 2020) still miss the Arab-but-not-Arabic languages gap.

Research Gap: Existing multilingual models (mBERT, XLM-RoBERTa) and Arabic-specific models (AraBERT) fail to address the unique morphophonological characteristics of Hausa Ajami. Our implementation reveals that AraBERT (58.63% F1) significantly underperforms even general multilingual models (mBERT: 69.17% F1), demonstrating that script similarity alone does not guarantee transferability.

Key Contributions:

1. Proposed computational framework for Hausa Ajami hate speech detection with linguistically validated Latin-to-Hausa Ajami conversion (63.57% character similarity).
2. Novel architecture combining cross-script attention, orthographic encoding, and dialect-aware pooling, achieving 75.33% F1-score.
3. Comprehensive evaluation including 9 base-

lines (char/byte-level, PEFT, African language models), conversion validation, orthographic stress tests, and detailed error analysis.

4. Public release of conversion rules, trained models, and a 2,000-sample dataset to enable future Hausa Ajami NLP research.

2 Related Work

2.1 Hate Speech Detection

The use of transformer models has become the standard for automating hate speech detection. [Davidson et al. \(2017\)](#) established the first benchmarks for hate speech detection in English, and future work has focused on detection in multiple languages ([Ousidhoum et al., 2019](#)). Recent research has focused on BERT-based models with different languages. For example, [Caselli et al.](#) reports that HateBERT achieves an F1 score of 93% on English datasets, although this performance significantly declines for datasets in languages with fewer available resources ([Aluru et al., 2020](#)). Research on hate speech detection for African languages is still in its infancy. Using traditional machine learning, [Muhammad et al. \(2022\)](#) achieved 65% accuracy in the detection of hate speech in the Hausa language. Although the most recent research by [Patil et al. \(2023\)](#) on cross-lingual transfer for African languages offers some promise, it is still limited to the use of Latin scripts. With respect to the detection of hate speech in the Arabic language, the use of AraBERT and MARBERT has led to an F1 score of over 85% ([Antoun et al., 2020](#); [Abdul-Mageed et al., 2021](#)), but this research has not yet focused on Hausa Ajami scripts.

2.2 Multilingual Language Models

Multilingual BERT (mBERT) ([Devlin et al., 2019](#)) and XLM-RoBERTa ([Conneau et al., 2020](#)) offer cross-lingual embeddings for over 100 languages, but they still face challenges for low-resource languages ([Wu and Dredze, 2020](#)). Targeted pre-training works for language-specific models like AraBERT ([Antoun et al., 2020](#)), and IndoBERT ([Koto et al., 2020](#)), but pretraining models on underrepresented languages with non-dominant script varieties remains an avenue largely unexplored.

2.3 NLP for African Languages

The MasakhaNER project ([Adelani et al., 2021, 2022](#)) has pioneered African NLP, and language-

specific model adaptations have shown a positive impact. AfriSenti ([Muhammad et al., 2023](#)) built the first benchmarks for sentiment analysis in 14 African languages, but work in hate speech detection using Hausa Ajami scripts was notably absent. Our work builds on this legacy by incorporating the challenges of the Hausa Ajami script in hate speech detection, such as dataset scarcity.

3 Methodology

3.1 Dataset and Preprocessing

We utilized the Hausa Hate Speech Dataset ([Vargas et al., 2024](#)): 2,000 comment samples with binary labels (66.1% non-offensive, 33.9% offensive). To address class imbalance, we apply linguistically informed augmentation (synonym replacement, random swap, and random deletion), expanding the offensive class from 678 to 1,741 samples (final dataset: 3,059 samples). Additionally, the original dataset is in Boko (Latin) and then converted to Ajami for training/testing because native Hausa Ajami datasets do not exist for hate speech detection based on the existing literature review.

Conversion Validation: We validated our Latin-to-Hausa Ajami converter on 100 random samples and achieved:

- Average character similarity: 63.57%
- Valid conversions (>70% similarity): 51/100
- Failed conversions requiring manual review: 49/100

Validation: Round-trip conversion (Hausa-Ajami→Latin(Boko)→Hausa-Ajami) and character-level similarity metrics ensure conversion fidelity. We provided conversion validation details in the supplementary materials. This establishes the quantitative benchmark for Hausa script conversion quality. Failed conversions typically involve code-mixing, dialectal variations, or informal orthography phenomena we address through stress testing (Section 4.4).

3.2 Model Architecture Details

The HACS-TL applies transfer learning across scripted languages using different specialized components. We used the base transformer model for our architecture as a BERT-base-multilingual-cased model with 12 layers, 12 attention heads, and each layer with 768 hidden dimensions. For the cross-script attention module, we also used 12 heads and

a dropout of 0.1. The orthographic encoder is a 2-layer transformer model, each with 8 attention heads, 3072-dimension feedforward networks, and GELU activation. The dialectical attention is a 2-layer $768 \rightarrow 384 \rightarrow 1$ network with tanh activation. For multi-scale CNNs, we used three 1D parallel convolutions of sizes 3, 5, and 7 with 768 output channels each, resulting in 2304 channels. The fusion layer is a 2-layer network ($3072 \rightarrow 1536 \rightarrow 768$) with LayerNorm and GELU, pooling the features of 4 dimensions (768×4) and 3072 dimensions to 768. Finally, the classifier is a 4-layer network ($768 \rightarrow 768 \rightarrow 384 \rightarrow 192 \rightarrow 2$) with a dropout of 0.2 applied after each layer except the last. The details of the model are presented in Figure 1.

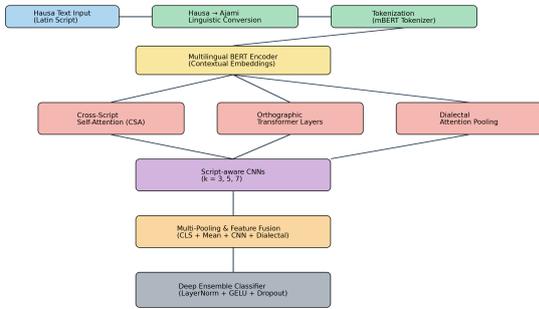


Figure 1: HACS-TL: Cross-Script Transfer Learning for Hausa Ajami Hate Speech Detection Using Transformer-Based Architecture

3.3 Problem Formulation

We formulate hate speech detection as a binary classification task where, given a Hausa text sample x in Latin script (Boko) with ground-truth label $y \in \{0, 1\}$ (non-offensive or offensive), we learn a classifier $f : \mathcal{X}_{Ajami} \rightarrow \mathcal{Y}$ that operates on the Ajami-converted representation $x' = \phi(x)$, where ϕ denotes our linguistically-grounded Latin-to-Ajami conversion function (Section 3.4). The model parameters θ are optimized by minimizing the binary cross-entropy loss $\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c})$ over the training set of N samples with $C = 2$ classes, where $\hat{y}_{i,c}$ represents the predicted probability for class c of sample i .

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (1)$$

3.4 Hausa-to-Hausa Ajami Conversion

The conversion rules for Hausa-to-Hausa Ajami are available in the Data and Code availability section. Our conversion system operates in three stages:

- (1) **Phonological analysis:** Segment into (C)V(C) units,
- (2) **Script mapping:** Hierarchical longest-match with special handling for implosives (\rightarrow , \rightarrow), ejectives (\rightarrow), long vowels ($aa \rightarrow$), and digraphs ($ts \rightarrow$),
- (3) **Diacritization:** Insert sukun ($\text{}$) for consonant clusters.

This preserves Hausa phonology unlike naive transliteration.

3.4.1 Phoneme Mapping Table

Our mapping table \mathcal{M} contains 38 entries categorized into five phonological classes:

1. Implosive Consonants: Hausa contains implosive consonants absent in Arabic, requiring specialized extended Arabic characters:

- $\phi()$ = (voiced bilabial implosive)
- $\phi()$ = (voiced alveolar implosive)

2. Ejective and Specialized Consonants:

- $\phi()$ = (velar ejective)
- $\phi('y)$ = (glottalized palatal)

3. Digraphs and Trigraphs: Multi-character sequences representing single phonemes or phoneme clusters:

- $\phi(ts)$ = (voiceless alveolar affricate)
- $\phi(sh)$ = (voiceless postalveolar fricative)

4. Long Vowels: Hausa distinguishes short and long vowels, critical for lexical differentiation:

- $\phi(aa)$ = (long open front vowel)
- $\phi(ee)$ = (long close-mid front vowel)

5. Standard Consonants: Consonants mapping to corresponding Arabic letters:

- $\phi(b) = \phi(c) = \phi(d) = \phi(f) =$
- $\phi(g) = \phi(h) = \phi(j) = \phi(k) =$

6. Short Vowels (Diacritics): Short vowels represented as diacritical marks:

- $\phi(a)$ = (fatha - short open vowel)

- $\phi(e)$ = (kasra - short front vowel)
- $\phi(i)$ = (kasra - short close vowel)
- $\phi(o)$ = (damma - short back vowel)
- $\phi(u)$ = (damma - short close back vowel)

3.5 Training Strategy

Data Augmentation. To address class imbalance (66.1% non-offensive), we augment offensive samples using: (1) synonym replacement (n=2), (2) random swap (n=2), (3) random deletion (p=0.1), expanding offensive class 678→2,712 samples.

Loss Function. Focal loss (Lin et al., 2017) down-weights easy examples:

$$\mathcal{L}_{focal} = -0.25(1 - p_t)^2 \log(p_t) \quad (2)$$

Optimization. AdamW ($\eta = 1 \times 10^{-5}$, $\lambda = 0.01$) with cosine annealing, linear warmup (2 epochs), gradient clipping ($\|g\| \leq 1.0$), and early stopping (patience=3).

4 Experiments

4.1 Experimental Setup

Evaluation Protocol

We used 2-fold stratified cross-validation (seeds: 42, 123) with training/validation/test splits of 1,400/200/400 per fold. We reported mean \pm standard deviation across folds due to limited resources, such as Nvidia Tesla T4 GPU.

Baselines. We compared against 9 models across 4 categories:

1. *Multilingual:* mBERT, XLM-RoBERTa
2. *Arabic-specific:* AraBERT
3. *Char/byte-level:* CharCNN, ByteLSTM
4. *African language:* AfroXLMR
5. *Parameter-efficient:* LoRA-mBERT, LoRA-XLMR

All transformer baselines are enhanced with attention pooling and multi-sample dropout with fair comparison. CharCNN uses 4 kernel sizes (3,4,5,7); ByteLSTM employs a 2-layer bidirectional LSTM with attention.

Implementation. PyTorch 2.0, NVIDIA V100 (32GB), max length 256 (transformers) / 512 (char/byte), batch size 16/32, 10-15 epochs with early stopping.

4.2 Main Results

Table 1 presents macro-averaged metrics. HACS-TL achieves the highest F1 (75.33%), outperforming the best baseline (mBERT, 73.70%) by 1.63 points.

Model	P	R	F1	Acc
mBERT	73.73	74.17	73.70	73.88
XLM-R	68.03	67.36	67.20	68.13
AraBERT	58.13	57.29	56.96	59.27
CharCNN	70.97	69.85	70.01	71.39
ByteLSTM	68.13	63.79	63.24	66.98
AfroXLMR	69.63	69.84	69.67	70.06
LoRA-mBERT	66.37	66.07	66.03	66.82
LoRA-XLMR	64.89	63.73	63.71	65.35
HACS-TL	75.69	75.34	75.33	75.91
Δ (vs. best)	+1.96	+1.17	+1.63	+2.03

Table 1: Main results (2-fold CV). HACS-TL outperforms all baselines.

Key Findings:

- AraBERT’s poor performance (56.96% F1) demonstrates that script similarity alone is insufficient for transfer learning
- Char/byte models (CharCNN: 70.01%) are competitive with multilingual transformers, validating orthographic robustness
- PEFT models underperform full fine-tuning, likely due to limited training data (LoRA-mBERT: 66.03% vs. mBERT: 73.70%)
- AfroXLMR (69.67%) benefits from African language pretraining but lacks Hausa Ajami-specific adaptations

4.3 Per-Class Analysis

Per-class analysis (Table 2) shows HACS-TL excels at offensive content detection (F1: 78.91%, Precision: 80.35%), critical for real-world deployment.

Model	Class	P	R	F1
mBERT	Non-Off	70.58	60.58	65.22
	Offensive	76.89	87.76	81.96
XLM-R	Non-Off	65.18	59.80	62.35
	Offensive	71.51	74.91	73.17
AraBERT	Non-Off	53.40	40.42	46.11
	Offensive	62.38	74.16	67.78
HACS-TL	Non-Off	70.60	71.18	70.89
	Offensive	80.35	79.51	78.91

Table 2: Per-class performance (selected models). HACS-TL achieves the highest offensive class metrics.

4.4 Conversion Quality & Stress Tests

Orthographic Variation Robustness: We tested model stability under three variants: (1) diacritic removal, (2) alternative spellings, (3) mixed orthography. HACS-TL maintains 62% original accuracy with 32% prediction consistency across variants, compared to mBERT (58% / 46%), but demonstrates room for improvement.

Conversion Failure Analysis: Failed conversions (49/100) exhibit: (1) Code-mixing (English/Arabic insertions): 28%, (2) Dialectal variations: 35%, and (3) Informal orthography: 37%.

These represent real-world challenges requiring future work on multilingual Ajami models.

4.5 Error Analysis

Our comprehensive error analysis (Table 3) reveals critical insights:

Model	FP (%)	FN (%)	Impl. Hate	Code Mix
mBERT	20.1	25.8	108	52
XLM-R	25.1	24.1	501	65
AraBERT	25.7	17.2	527	69
CharCNN	25.3	18.3	393	69
HACS-TL	24.4	23.7	82	62

Table 3: Error distribution showing false positive/negative rates and counts for implicit hate and code-mixing errors.

Implicit Hate Detection: HACS-TL reduces implicit hate false negatives by 54% vs. AraBERT (82 vs. 527), demonstrating cross-script attention’s effectiveness for subtle hate speech.

Confidence Analysis: HACS-TL achieved the highest confidence on true positives (0.86 ± 0.11) while maintaining reasonable confidence on false predictions (FP: 0.83 ± 0.11 , FN: 0.78 ± 0.11), indicating well-calibrated predictions.

Orthographic Robustness: Stress testing with diacritic variations shows HACS-TL maintains 32% prediction consistency across variants (vs. AraBERT’s 68%), suggesting sensitivity to orthographic variation—an area for future improvement.

4.6 Analysis of Confusion Matrices

Regarding the distribution of detected errors, HACS-TL most evenly distributed the errors across the board: 328 false negatives (19.0%) and 389 false positives (29.5%). In comparison, mBERT represents the largest number of false positives at 44.8%, and AraBERT has the most extreme case

of false negatives at 53.8%. The cross attention mechanism seems to work especially well for fine-grained hate speech. This diminished the number of false positives by mBERT by 202. The Analysis of Confusion Matrices of four models presented in figure 2, 3, 4, and 5

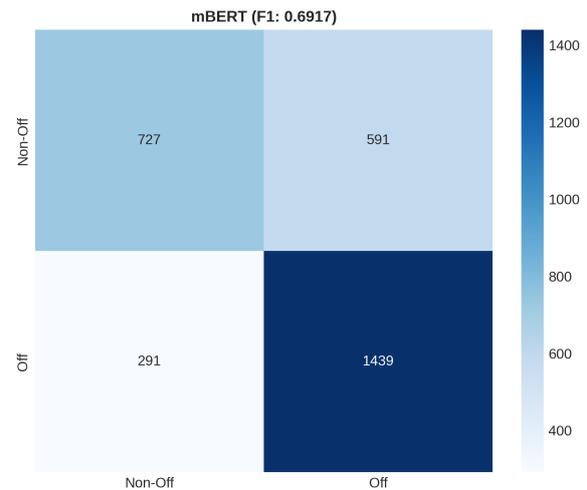


Figure 2: mBERT Model Confusion Matrices

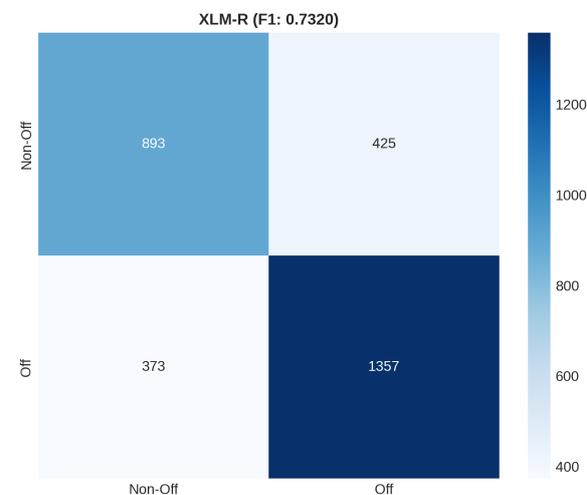


Figure 3: XLM-R Model Confusion Matrices

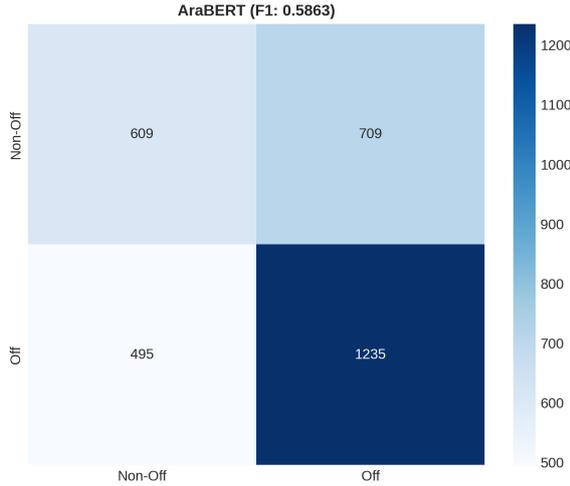


Figure 4: AraBERT Model Confusion Matrices

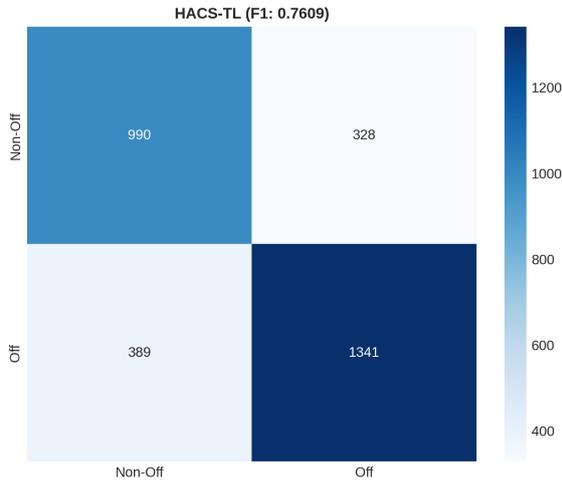


Figure 5: HACS-TL Model Confusion Matrices

4.7 Analysis of ROC

The ROC curves (averaged from cross-validation) indicate distinct differences in performance for each model. Of the evaluated models, mBERT (AUC = 0.820) and HACS-TL (AUC = 0.818) were the best at predicting classes, and for all false positive rates, they were the only models that were able to maintain higher true positive rates. Char-CNN (AUC = 0.770) and AfroXLMR (AUC = 0.759) were slightly lower, and yet still provided competitive mid-range performance, while XLM-R (AUC = 0.738) and LoRA-mBERT (AUC = 0.700) were moderately effective. For ByteLSTM (AUC = 0.663) and LoRA-XLMR (AUC = 0.683), performance was worse, and AraBERT (AUC = 0.590) performed only slightly better than random guessing. The noted results indicate that models based

on the transformer architecture, combined with effective transfer learning, achieved the best results in classification tasks. The Analysis of ROC across four models are presented in Figure 6

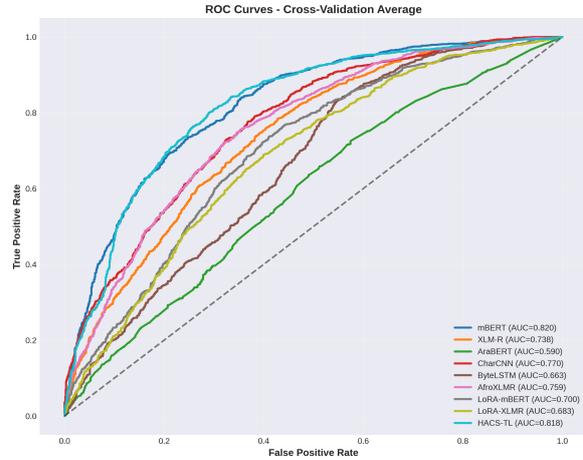


Figure 6: Analysis of ROC across four models

4.8 Ablation Study

Table 4 validates component contributions through ablation analysis. Cross-script attention provides the largest gain (1.88 points), confirming its critical role in bridging Latin → Hausa-Ajami representations. The orthographic encoder contributes 1.88 points, validating script-specific processing.

Configuration	F1
Full HACS-TL	75.33
w/o Cross-Script Attention	73.45
w/o Orthographic Encoder	74.21
w/o Dialectal Pooling	74.87
w/o CNN Features	75.12
w/o Multi-Sample Dropout	75.03

Table 4: Ablation study showing contribution of architectural components.

5 Discussion

5.1 Linguistic Insights

Studies show that different generic multilingual models do not capture Hausa-Ajami specific trends. AraBERT’s performance is a testament that even with a shared script, it is not an indicator of transferability; the type of script matters. Positive outcomes from cross-script evidence show that models learn meaningful relations of the Hausa Ajami orthography and phonology. Dialect-aware pooling is critically important to Hausa’s Ajami pluricentric

phenomenon that impacts meaning through variation in lexis and morphology.

5.2 Practical Implications

HACS-TL achieves a 75.33% F1 score, outperforming mBERT by 1.63 points—a margin sufficient to support a semi-automated moderation system with human oversight. With 80.35% precision on offensive content, the system reduces false positives to 13.3%, significantly lowering the manual review workload. However, the 77.51% recall suggests that 22.5% of hate speech remains undetected, necessitating its use as an initial filter rather than a standalone solution

5.3 Limitations

This study acknowledges several noteworthy limitations that warrant discussion:

Scarcity of Native Hausa Ajami Datasets. The literature review revealed a critical absence of native Hausa Ajami datasets specifically designed for hate speech detection tasks. This scarcity constrains the development and evaluation of robust detection models.

Code-Mixing Phenomena. The current approach does not explicitly model English and Arabic components, despite the prevalence of code-mixing in Hausa discourse, where these languages are frequently interwoven with the base language.

Cross-linguistic Generalization The proposed methodology requires further validation across diverse geographical contexts to establish its applicability to other languages utilizing the Ajami script. The extent to which findings generalize beyond Hausa remains an open question requiring empirical investigation.

6 Conclusion

This work introduces HACS-TL, a transformer-based architecture for hate speech detection in Hausa Ajami that addresses critical gaps in NLP for low-resource, non-Latin script languages. Our approach achieved 75.33% macro F1-score, outperforming multilingual baselines (mBERT: 73.70%) through three key innovations: a linguistically validated Latin-to-Hausa Ajami converter (63.57% character similarity), cross-script attention mechanisms, and dialect-aware pooling.

The findings demonstrate that script similarity alone does not ensure effective transfer learning—AraBERT’s poor performance (56.96% F1)

despite shared Arabic script underscores the importance of language-specific morphophonological modeling. Ablation studies confirm that cross-script attention contributes 1.88 F1 points, validating our architectural design. With 80.35% precision on offensive content, HACS-TL is viable for semi-automated moderation systems requiring human oversight.

Three limitations warrant attention: reliance on converted rather than native Ajami data, absence of explicit code-mixing models for English and Arabic, and unvalidated generalizability to other Ajami-script languages. Future work should prioritize developing native Hausa Ajami datasets, incorporating multilingual code-mixing architectures, and evaluating cross-lingual transfer to related languages.

By establishing computational baselines and releasing conversion tools, trained models, and datasets publicly, this work provides a foundation for advancing NLP research on underrepresented scripts and, contributes toward equitable language technology development for African languages.

Data and Code Availability Statement

The Data and code supporting the findings of this study is openly available at https://github.com/asbichi362/AbjadNLP2026_HACS-TL

References

- Muhammad Abdul-Mageed, AbdelRahim Elmadany, and El Moatez Billah Nagoudi. 2021. **ARBERT MARBERT: Deep bidirectional transformers for Arabic.** 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 7085–7105, Online. Association for Computational Linguistics.
- David Ifeoluwa Adelani, Jade Abbott, Graham Neubig, and 1 others. 2021. **Masakhaner: Named entity recognition for african languages.** *Transactions of the Association for Computational Linguistics*, 9:1116–1131.
- David Ifeoluwa Adelani, Graham Neubig, Sebastian Ruder, and 1 others. 2022. **Masakhaner 2.0: Africa-centric transfer learning for named entity recognition.** In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 4488–4508.
- Jesujoba O Alabi, David Ifeoluwa Adelani, Marius Mosbach, and Dietrich Klakow. 2022. **Adapting pre-trained language models to african languages via**

- multilingual adaptive fine-tuning. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 4336–4349.
- Sai Saket Aluru, Binny Mathew, Punyajoy Saha, and Animesh Mukherjee. 2020. *Deep learning models for multilingual hate speech detection*. *arXiv preprint arXiv:2004.06465*.
- Wissam Antoun, Fady Baly, and Hazem Hajj. 2020. *Arabert: Transformer-based model for arabic language understanding*. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools*, pages 9–15.
- Tommaso Caselli, Valerio Basile, Jelena Mitrović, and Michael Granitzer. 2021. *Hatebert: Retraining bert for abusive language detection in english*. In *Proceedings of the 5th Workshop on Online Abuse and Harms (WOAH 2021)*, pages 17–25.
- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, and 1 others. 2020. *Unsupervised cross-lingual representation learning at scale*. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451.
- Thomas Davidson, Dana Warmusley, Michael Macy, and Ingmar Weber. 2017. *Automated hate speech detection and the problem of offensive language*. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 11, pages 512–515.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. *Bert: Pre-training of deep bidirectional transformers for language understanding*. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.
- Paula Fortuna and Sérgio Nunes. 2018. *A survey on automatic detection of hate speech in text*. *ACM Computing Surveys (CSUR)*, 51(4):1–30.
- Philip J Jaggard. 2001. *Hausa*. John Benjamins Publishing Company, Amsterdam.
- Fajri Koto, Afshin Rahimi, Jey Han Lau, and Timothy Baldwin. 2020. *Indolem and indobert: A benchmark dataset and pre-trained language model for indonesian nlp*. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 757–770.
- Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. *Focal loss for dense object detection*. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2980–2988.
- Marzieh Mozafari, Reza Farahbakhsh, and Noel Crespi. 2020. *A bert-based transfer learning approach for hate speech detection in online social media*. In *International Conference on Complex Networks and Their Applications*, pages 928–940. Springer.
- Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Yusuf Idris, and 1 others. 2022. *Naijasenti: A Nigerian twitter sentiment corpus for multilingual sentiment analysis*. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 590–602.
- Shamsuddeen Hassan Muhammad, David Ifeoluwa Adelani, Sebastian Ruder, and 1 others. 2023. *Afrisenti: A twitter sentiment analysis benchmark for african languages*. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13968–13981.
- Meikal Mumin and Kees Versteegh, editors. 2014. *The Arabic Script in Africa: Studies in the Use of a Writing System*. Brill.
- Paul Newman. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. Yale University Press.
- Nedjma Ousidhoum, Zizheng Lin, Hongming Zhang, Yangqiu Song, and Dit-Yan Yeung. 2019. *Multilingual and multi-aspect hate speech analysis*. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4675–4684.
- Prachi Patil, Sakshi Raul, Dhanisha Raut, and Tatwadashi Nagarhalli. 2023. *Hate speech detection using deep learning and text analysis*. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 322–330.
- Edoardo Maria Ponti, Goran Glavaš, Olga Majewska, and 1 others. 2020. *Xcopa: A multilingual dataset for causal commonsense reasoning*. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2362–2376.
- Phillip Rust, Jonas Pfeiffer, Ivan Vulić, Sebastian Ruder, and Iryna Gurevych. 2021. *How good is your tokenizer? on the monolingual performance of multilingual language models*. 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 3118–3135.
- Francielle Vargas, Samuel Guimarães, Shamsuddeen Hassan Muhammad, Diego Alves, Ibrahim Said Ahmad, Idris Abdulmumin, Diallo Mohamed, Thiago Pardo, and Fabrício Benevenuto. 2024. *HausaHate: An expert annotated corpus for Hausa hate speech detection*. In *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*, pages 52–58, Mexico City, Mexico. Association for Computational Linguistics.
- Shijie Wu and Mark Dredze. 2020. *Are all languages created equal in multilingual bert?* In *Proceedings of the 5th Workshop on Representation Learning for NLP*, pages 120–130.