

CogniCIC at HOMO-LAT 2025: Fine-tuning and Semantic Similarity Approaches for Polarity Detection in Spanish Texts

Cesar Macias^{1,†}, Miguel Soto^{1,*,†}, Tania Alcántara^{1,†}, Omar García^{1,†},
José Eduardo Valdez-Rodríguez¹ and Edgardo Felipe-Riveron¹

¹Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional (IPN), México

Abstract

Hate speech against the LGBTQ+ community in digital environments represents a growing challenge that requires automatic solutions that are able to adapt to the dialectal and cultural diversity of Spanish. This paper presents two methodologies based on deep learning models for polarity classification of messages on Reddit targeting the LGBTQ+ community. The first methodology (Track 1) addresses a classification of polarity (positive, negative or neutral) using a shared corpus of tweets and posts on Reddit generated in Mexico, Colombia, Chile and Argentina. The second methodology (Track 2) focuses on a cross-dialect classification, training on dialects from the centre (Mexico, Colombia, Chile, Argentina) and evaluating on dialects from the rest of Latin America (Bolivia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Uruguay and Venezuela). Our best results reach a *F1-score* of 0.4350 and 0.4388 for Tracks 1 and 2, respectively, evidencing the complexity of dialectal variants and the need to incorporate adapted preprocessing, more data and additional techniques to improve nuance detection.

Keywords

hate speech, LGBTQ+ community, multi-dialect classification, dialect cross-classification, deep learning

1. Introduction

Hate speech directed at the LGBTQ+ community in digital environments has gained significant relevance in recent years due to its social and psychological implications. The proliferation of online platforms facilitates the spread of discriminatory messages, which can foster violence and social exclusion of people who are part of the community. For this reason, automatic detection of hate speech against the LGBTQ+ community is essential to help build artificial intelligence models or architectures to implement an effective moderation system, which is able to identify and mitigate hate speech across multiple languages and domains.

Internationally, competitions such as LT-EDI (Language Technology for Equality, Diversity, Inclusion) have promoted the development of Natural Language Processing (NLP) approaches for homophobia and transphobia classification in social network comments, demonstrating that models based on transformer-like architectures (e.g., mBERT [1], XLM-R [2]) far outperform traditional methods in multilingual tasks [3, 4]. In the Spanish-speaking domain, the HOMO-MEX competition organized at IberLEF (Iberian Languages Evaluation Forum) 2023 provided the first annotated corpus of tweets in Mexican Spanish

IberLEF 2025, September 2025, Zaragoza, Spain

*Corresponding author.

†These authors contributed equally.

✉ cmaciass2021@cic.ipn.mx (C. Macias); msotoh2021@cic.ipn.mx (M. Soto); talcantaram2020@cic.ipn.mx (T. Alcántara); ogarciav2024@cic.ipn.mx (O. García); jvaldezr2018@cic.ipn.mx (J.E. Valdez-Rodríguez); edgardo@cic.ipn.mx (E. Felipe-Riveron)

🌐 <https://www.linkedin.com/in/cesar-macias-sanchez-829919367/> (C. Macias);

<https://www.linkedin.com/in/miguel-angel-soto-hernandez/> (M. Soto); <https://www.linkedin.com/in/talcantaram/>

(T. Alcántara); <https://www.linkedin.com/in/omar-garcia-vazquez-093128219/> (O. García);

<https://www.linkedin.com/in/jose-eduardo-valdez-rodriguez-38720b117/> (J.E. Valdez-Rodríguez)

🆔 0009-0005-1708-5359 (C. Macias); 0009-0006-4619-9352 (M. Soto); 0009-0001-4391-6225 (T. Alcántara);

0009-0002-1205-1166 (O. García); 0000-0002-4572-5713 (J.E. Valdez-Rodríguez); 0000-0002-9828-3568 (E. Felipe-Riveron)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

for LGBT+phobia detection. For 2024, it extended the challenge to more complex domains, such as song lyrics. These efforts have shown that, while Spanish transformers (BETO [5] or mBERT [1]) achieve F1-macro scores above 0.88 in binary detection, challenges remain in capturing nuances of implicit hatred and adapting to particular dialectal variants and cultural contexts.

While existing competencies have already laid a solid foundation, the linguistic and sociocultural diversity of Spanish-speaking countries demands specific resources that take into account regional variations of Spanish. The HOMO-LAT 2025 [6] is a shared task of the shared evaluation campaign IberLEF 2025 [7]. This task aims to extend the analysis to the Spanish of several Latin American countries on platforms such as Reddit, focusing on the polarity of LGBTQ+ mentions (positive, negative or neutral) and addressing the challenges associated with local contexts. In this article, we propose the design and implementation of two models for the automatic classification of hate speech against the LGBTQ+ community, both based on deep learning techniques and preprocessing adapted to Spanish characteristics, with a view to proposing a model capable of strengthening online moderation and inclusion.

2. State of the Art

In recent years, the automatic detection of anti-LGBTQ+ hate speech has evolved significantly, driven by the availability of new annotated resources and the refinement of deep language models. Earlier in the decade, research focused on the creation of multilingual corpora and descriptive analysis of hate phenomena in different linguistic contexts. In 2021, Chakravarthi et al. introduced a pioneering corpus of YouTube comments in English and Dravidian languages (Tamil, Malayalam, and Tamil-English mixed text) annotated for homophobia and transphobia, serving as the basis for the first LT-EDI competition in 2022 [8]. In the same year, Plaza-del Arco et al. comprehensively evaluated various pre-training models (BETO [5], mBERT [1], XLM-R [2]) and traditional architectures (CNN [9], LSTM [10]) on a corpus of Spanish tweets to detect hate speech, including homophobic insults, demonstrating the superiority of Spanish monolingual models over multilingual and traditional approaches [11]. Similarly, Hudhayri, in 2021, documented, from a qualitative perspective, forms of linguistic harassment against LGBTQ+ people on Arabic Twitter (now X), identifying verbal and visual patterns of homophobic aggression that highlight the need for Arabic-specific PLN tools [12].

During 2022, the ranking competitions were consolidated with the second LT-EDI Workshop. Chakravarthi et al. described the first edition of the Shared Task in LT-EDI 2022, where around 50 teams competed in detecting homophobia and transphobia in social media comments in English, Tamil, and mixed code Tamil-English. The results showed that the multilingual transformer-based models outperformed traditional classifiers by far, achieving F1-macro scores close to 0.80 in the English and Tamil versions [3]. In the same year, Karayıgıt et al. explored the Turkish language, building the HATC corpus to detect homophobic and generic hate speech on Instagram. By comparing mBERT [1] with neural network architectures and classical classifiers (SVM [13], Random Forest [14], GRU [15]), they showed that the transformer-based architecture (mBERT [1]) achieved an average F1 of 0.90, accurately capturing the subtle homophobic expressions in Turkish [16].

In 2023, attention shifted to Mexican Spanish resources and the refinement of shared tasks. Vázquez et al. presented HOMO-MEX, the first annotated corpus of Mexican Spanish tweets for LGBT+phobia detection, with approximately 7,000 tweets tagged in binary and multi-label categories, serving as an initial benchmark where again, BETO [5] demonstrated outstanding performance with F1 of 0.74 in binary classification and 0.85 in fine-granular classification, leaving as evidence the complexity of the homophobic language [17]. In parallel, Bel-Enguix et al. described the 2023 edition of the shared task HOMO-MEX in IberLEF, where with the use of similar models, participants managed to obtain superior results, emphasizing the importance of careful preprocessing in Spanish [18]. In the same year, Chakravarthi et al. conducted the second edition of LT-EDI 2023, expanding the languages to five (English, Spanish, Tamil, Hindi and Malayalam) and adding a sub-task with seven hate classes. The use of transformer-based models again outperformed classical classification models, resulting in

F1s of 0.90 in English and 0.85 in Spanish and Tamil, although performance was lower in Hindi and Malayalam due to the paucity of data [19]. Kumaresan et al. complemented these efforts by presenting a high-quality corpus for detecting homophobia and transphobia in Hindi and Malayalam in 2023, where they trained a model with both languages achieving an F1-macro close to 0.80 in Hindi and close to 0.75 in Malayalam, pointing out the benefit of sharing information between related languages [20].

In 2024, the challenges diversified into specialised domains and socio-temporal analyses. For the 2024 edition of HOMO-MEX at IberLEF, Gómez-Adorno et al. introduced three sub-tasks: general hate detection in tweets, fine-granular classification by type of phobia and detection of homophobic content in song lyrics. The winning proposals achieved Macro-F1 close to 0.91 on tweets and up to 0.97 on fine classification, but the music lyrics task proved more difficult reaching an F1 of 0.58, demonstrating the need for domain-specific models and additional data [21]. Similarly, Chakravarthi et al. presented the third edition of LT-EDI at EACL 2024, extending the competition to ten languages (including South Indian languages such as Telugu, Kannada, Gujarati, Marathi and Tulu). Systems based on XLM-R [2] and specific monolingual models achieved F1s between 0.80 and 0.95, demonstrating that joint learning of multilingual data benefits under-resourced languages, although challenges remain in capturing cultural nuances [4]. Complementarily, Andersen et al. analysed the sociolinguistic evolution of LGBT+ related terms on Mexican Twitter between 2011 and 2021, using context vectors and sentiment analysis to demonstrate changes in the semantic polarity of slurs and neutral terms, highlighting that detection models require constant temporal adjustments to adapt to language evolution [22]. Finally, this year, McGiff and Nikolov explored the difference in performance between classical and BERT [1] models for detecting homophobia on Twitter (now X), highlighting the relevance of capturing context for identifying implicit homophobia [23].

Finally, in 2025, Leoni Santos et al. applied deep learning and explainability techniques to a specific domain: tweets homologous to the European Football Championship, where homophobia is prevalent. They introduced H-DICT, a specialized dictionary for filtering and annotating homophobic tweets, and compared five variants of BERT [1], where RoBERTa Offensive [24] obtained the best result with an F1 close to 0.89, and using Integrated Gradients they identified the words guiding the detection, concluding that adapting models to the domain and using specialized lexicons substantially improves the detection of anti-LGBT+ hate in sporting contexts [25]. The HOMO-LAT [6] task is also announced at IberLEF 2025, extending the analysis to Reddit throughout Latin America, focusing on the polarity of mentions of LGBTQ+ terms (positive, negative or neutral) and addressing dialectal variants and local contexts, reflecting the constant evolution and expansion of research in this field.

3. Methodology

A brief description of the provided datasets for the competition is given in this section; furthermore, a detailed description of the method selected to resolve the problem is provided.

3.1. Tracks description

3.1.1. Track 1: Multi-dialect polarity detection track (multi-class)

For this track, the organizers proposed a dataset composed of Reddit posts that contained LGBTQ+ keywords. The keywords were defined by the organizers in a lexicon format; some of those keywords are: {trans, LGBT, queer, ...}. The objective of this task was to indicate the polarity {positive, negative, neutral} of the posts towards the keyword. The posts collected in this dataset were written in Spanish dialects from Mexico, Colombia, Chile, and Argentina. Training and testing datasets contained posts from every dialect; a detailed description of the datasets is presented in Section 3.2.

3.1.2. Track 2: Cross-dialect polarity detection (multi-labeled)

The training dataset for this track is the same as in Track 1. The difference from Task 1 was that the Spanish dialects contained in the testing corpus were not the same as in the training set. The posts

contained in the testing set are written in Spanish dialects from Bolivia, Costa Rica, Cuba, the Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Uruguay, and Venezuela. The goal of this task was to predict the polarity of the post, as in Task 1. A more detailed description of the training and testing set is presented in Section 3.2.

3.2. Corpus description

The corpus for this Shared Task is composed of three subsets {Train (shared by both tasks), Task 1 Test, Task 2 Test}. The distribution of the posts by country of the Train and Task 1 Test corpus is shown in Figure 1a. The distribution of the posts by country of the Task 2 Test corpus is shown in Figure 1b. Finally, the label distribution by country of the train corpus is shown in Figure 2

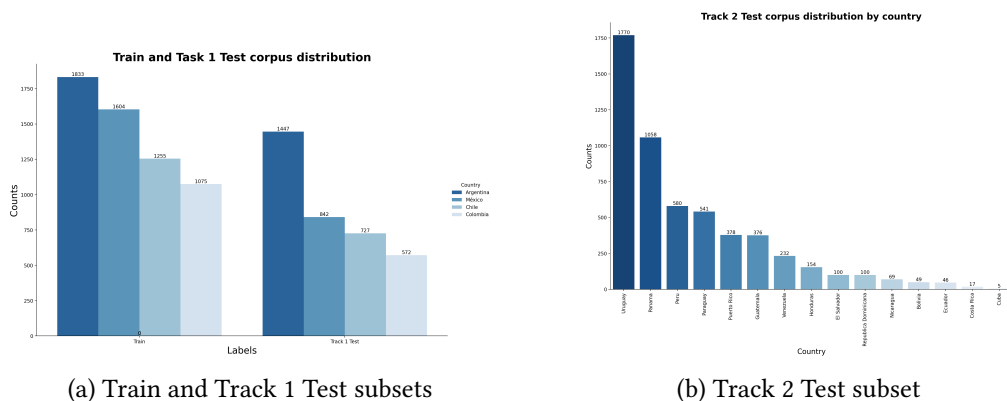


Figure 1: Corpus posts distribution by country for every subset

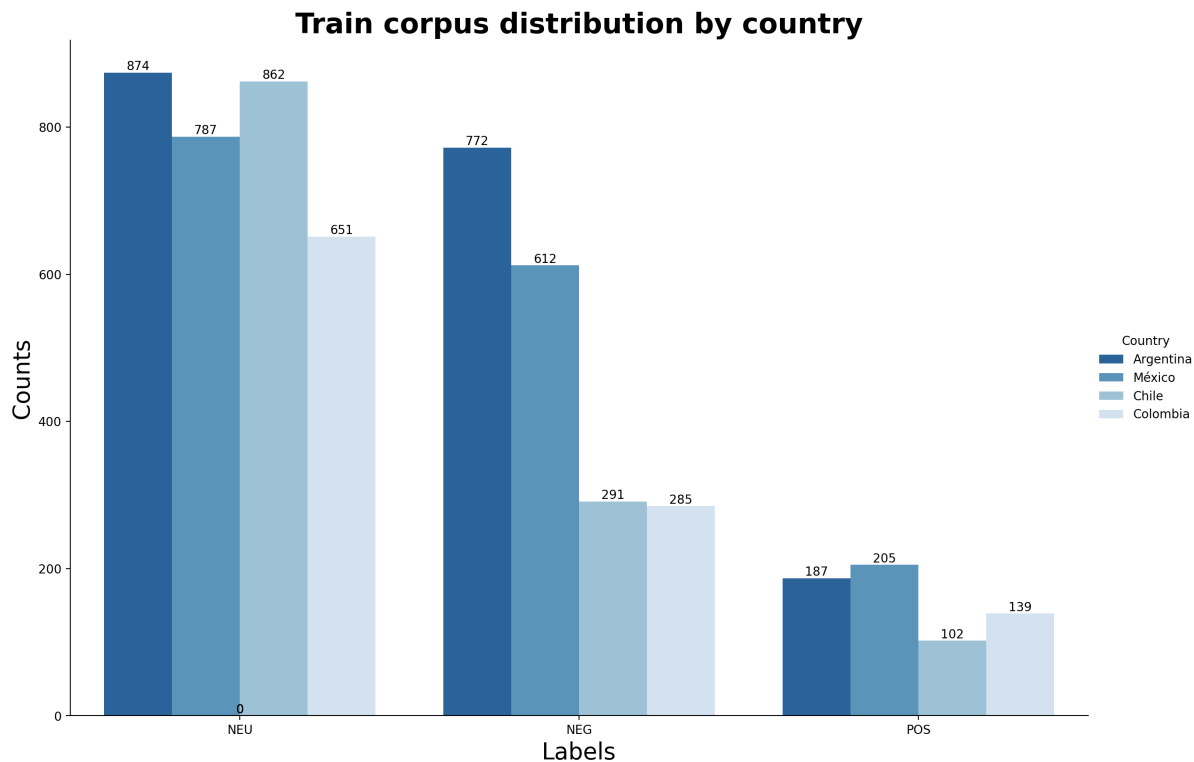


Figure 2: Train corpus label distribution by country

3.3. Model Specifications

3.3.1. Model one

Our approach for this shared task was to fine-tune a BERT [1] model pretrained in a large Spanish corpus, better known as BETO [5]. To optimize the model performance, we have configured several key specifications. Just one model was trained due to the existence of a sole training corpus; this model was trained for 25 epochs, we have selected the best model amongst all (tested over validation split against F1-score) to make the final predictions, a batch size of 64, the learning rate was set to 5×10^{-6} and the epsilon value was set to 1×10^{-8} , the validation size was set to be the 20% of the training set and finally the environment random generation seeds were set to 42. This model was used to make predictions over the testing sets of Track 1 and Track 2.

3.3.2. Model two

For subtasks 1 and 2, we used the BETO model [5], a Spanish pre-trained variant of BERT. The model was fine-tuned using the training dataset provided. Afterward, the development dataset was split into two equal parts: the first was used to monitor the fine-tuning process and select the best performing weights, while the second was used to evaluate the classification system.

Once the model was adjusted, the semantic representations (embeddings) for each text were generated. Instead of using the vector associated with the [CLS] token, a *mean pooling* strategy was applied over the last attention layers. This technique has been shown to be more robust for sentence-level classification tasks, as it better captures the overall semantic meaning of the text.

The generated vectors were stored in a vector database using OpenSearch [26]. For classification, the test data set was processed in the same way, and their embeddings were used to perform a semantic similarity search using the *k*-Nearest Neighbors (kNN) algorithm, with cosine similarity as the main distance metric. In addition to semantic search, a filtering step was applied based on country and keyword (LGBT + segment). This allowed the system to restrict the search to contextually similar vectors, considering that idiomatic expressions and community references can vary significantly between Spanish-speaking countries.

4. Experiments and results

Here we present the final ranking results obtained for the predictions made with the fine-tuned BETO model. The predictions were made with the best model tested over validation; the best performing model obtained an F1-score of 0.5381 and this was used to make the final predictions. The results obtained with the best validation model for Tracks 1 and 2 are shown in Table 1, and Table 2 respectively.

Table 1

Track 1 submissions. Final prediction scores

Rank	Team	Score
1	pepeolivert	0.5296
2	mary_toapanta	0.5261
3	rogeliocampos	0.5137
4	mcardoso94	0.4661
5	taiaille	0.4360
6	cmsdev	0.4300
7	ymlopez	0.2592

Table 2

Track 2 submissions. Final prediction scores

Rank	Team	Score
1	pepeolivert	0.5086
2	mary_toapanta	0.4803
3	rogeliocampos	0.4639
4	taiaille	0.4388
5	cmsdev	0.4054
6	mcardoso94	0.3622

5. Discussion

The results obtained in both tracks suggest that, although the fine-tuned BETO model achieves competitive validation scores, the generalisation on the test set reveals a significant gap against the best-ranked systems (Track 1: 0.5296 vs. 0.4360 reported and Track 2: 0.5086 vs. 0.4388 obtained). This difference could be explained mainly by the difference between training and test data, i.e. in Track 1, the test dialects (Mexico, Colombia, Chile and Argentina) were part of the training corpus, which favoured the adaptation of our models to the polarity of the messages. On the other hand, in Track 2, when testing on unseen dialects (e.g. Bolivia, Cuba, Ecuador or Venezuela), performance decreased by more than 10 percentage points in F1. This tells us that the contextual representation of our models loses robustness to idiomatic expressions and localisms not contemplated in the training.

The dialectal disparity highlights the importance of adapted pre-processing and strategies that mitigate linguistic variability. That is, the second methodology based on mean pooling of BETO embeddings and kNN classification with country and keyword filtering could help to retrieve semantically close examples despite lexical differences. However, its effectiveness depends directly on the diversity and quality of the reference corpus, so without a large repository of local expressions, it risks generating false positives or negatives. Furthermore, the variety in the distribution of tags by country and the presence of spelling variants and colloquialisms on Reddit (abbreviations, emojis, misspellings, irony) require incorporation through normalisation techniques and perhaps data augmentation (e.g., spelling correction and hashtag-sensitive tokenization) to improve the model’s ability to capture specific nuances and tones of hate speech.

For future research, it is essential to explore class balancing methods (oversampling of minority instances or weighted loss) and to expand BETO’s vocabulary by vocabulary expansion to include idioms and localisms. Furthermore, the integration of explainability modules (e.g., integrated gradients or *LIME*) would facilitate the identification of biases in polarity prediction and guide the construction of specialised lexicons, as demonstrated by the use of *H-DICT* in sports contexts. Taken together, these implementations would aim to strengthen the automatic detection of hate speech against the LGBTQ+ community in Latin America and the interpretation of pragmatic nuances in short social media texts.

Acknowledgments

The authors wish to thank the support of the Instituto Politécnico Nacional (COFAA, SIP-IPN, Research grants SIP-20251352, SIP-), and the Mexican Government (SECIHTI, SNI).

Declaration on Generative AI

During the preparation of this work, the authors used Writefull’s model in order to: Grammar and spelling check. Further, the authors used DeepL Translator in order to: Translate texts from Spanish to English.

References

- [1] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2019. URL: <https://arxiv.org/abs/1810.04805>. arXiv:1810.04805.
- [2] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, CoRR abs/1911.02116 (2019). URL: <http://arxiv.org/abs/1911.02116>. arXiv:1911.02116.
- [3] B. R. Chakravarthi, R. Priyadharshini, T. Durairaj, J. P. McCrae, P. Buitelaar, 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 (LT-EDI), 2022, pp. 1–6.
- [4] B. R. Chakravarthi, P. K. Kumaresan, R. Priyadharshini, P. Buitelaar, A. Hegde, H. Shashirekha, S. Rajiakodi, M. Á. García, S. M. Jiménez-Zafra, J. García-Díaz, R. Valencia-García, K. K. Ponnusamy, P. Shetty, D. García-Baena, Overview of third shared task on homophobia and transphobia detection in social media comments, in: Proceedings of the Fourth Workshop on LT-EDI (EACL), 2024, pp. 1–10.
- [5] J. Cañete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang, J. Pérez, Spanish pre-trained bert model and evaluation data, in: PML4DC at ICLR 2020, 2020.
- [6] G. Bel-Enguix, H. Gómez-Adorno, S. Ojeda-Trueba, G. Sierra, J. Barco, E. Lee, J. Dunstan, R. Manrique, Overview of HOMO-LAT at IberLEF 2025: Human-centric polarity detection in Online Messages Oriented to the Latin American-speaking lgbtq+ populaTion, Procesamiento del lenguaje natural 75 (2025) –.
- [7] J. Á. González-Barba, L. Chiruzzo, S. M. Jiménez-Zafra, Overview of IberLEF 2025: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2025), co-located with the 41st Conference of the Spanish Society for Natural Language Processing (SEPLN 2025), CEUR-WS. org, 2025.
- [8] B. R. Chakravarthi, R. Priyadharshini, R. Ponnusamy, P. K. Kumaresan, et al., Dataset for identification of homophobia and transphobia in multilingual youtube comments, arXiv preprint arXiv:2109.00227 (2021).
- [9] Y. Lecun, Y. Bengio, Convolutional Networks for Images, Speech and Time Series, The MIT Press, 1995, pp. 255–258.
- [10] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (1997) 1735–1780.
- [11] F. M. Plaza-del Arco, M. D. Molina-González, L. A. Ureña-López, M. T. Martín-Valdivia, Comparing pre-trained language models for spanish hate speech detection, Expert Systems with Applications 166 (2021) 114120. doi:10.1016/j.eswa.2020.114120.
- [12] K. Hudhayri, Linguistic harassment against arab lgbs on cyberspace, International Journal of English Linguistics 11 (2021) 58–75. doi:10.5539/ijel.v11n4p58.
- [13] C. Cortes, V. Vapnik, Support-vector networks, Machine learning 20 (1995) 273–297.
- [14] T. K. Ho, Random decision forests, in: Proceedings of 3rd international conference on document analysis and recognition, volume 1, IEEE, 1995, pp. 278–282.
- [15] J. Chung, Ç. Gülçehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, CoRR abs/1412.3555 (2014). URL: <http://arxiv.org/abs/1412.3555>. arXiv:1412.3555.
- [16] H. Karayigit, A. Akdağlı, İnan Aci, Homophobic and hate speech detection using multilingual-bert model on turkish social media, Information Technology and Control 51 (2022) 346–363. doi:10.5755/j01.itc.51.2.29988.
- [17] J. Vásquez, S. Andersen, G. Bel-Enguix, H. Gómez-Adorno, S.-L. Ojeda-Trueba, HOMO-MEX: A Mexican Spanish annotated corpus for LGBT+phobia detection on Twitter, in: Proceedings of the 7th Workshop on Online Abuse and Harms (WOAH), Association for Computational Linguistics, 2023, pp. 202–214.
- [18] G. Bel-Enguix, H. Gómez-Adorno, G. Sierra, J. Vásquez, S. T. Andersen, S.-L. Ojeda-Trueba, Overview of homo-mex at iberlef 2023: Hate speech detection in online messages directed towards

the mexican spanish speaking lgbtq+ population, *Procesamiento del Lenguaje Natural* 71 (2023) 361–370.

- [19] B. R. Chakravarthi, R. Ponnusamy, S. Malliga, P. Buitelaar, M. Á. García-Cumbreras, S. M. Jiménez-Zafra, J. A. García-Díaz, R. Valencia-García, N. Jindal, Overview of second shared task on homophobia and transphobia detection in social media comments, in: *Proceedings of the Third Workshop on LT-EDI*, 2023, pp. 1–10.
- [20] P. K. Kumaresan, R. Ponnusamy, R. Priyadharshini, P. Buitelaar, B. R. Chakravarthi, Homophobia and transphobia detection for low-resourced languages in social media comments, *Natural Language Processing (Elsevier) Journal* 5 (2023) 100041. doi:10.1016/j.nlp.2023.100041.
- [21] H. Gómez-Adorno, G. Bel-Enguix, G. Sierra, J. Vázquez, S. T. Andersen, S.-L. Ojeda-Trueba, Overview of homo-mex at iberlef 2024: Hate speech detection towards the mexican spanish speaking lgbt+ population, *Procesamiento del Lenguaje Natural* 73 (2024) 393–405.
- [22] S. T. Andersen, S.-L. Ojeda-Trueba, J. Vázquez, G. Bel-Enguix, The mexican gayze: A computational analysis of the attitudes towards the lgbt+ population in mexico on social media across a decade, in: *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH)*, Association for Computational Linguistics, 2024, pp. 178–200.
- [23] J. McGiff, N. S. Nikolov, Bridging the gap in online hate speech detection: A comparative analysis of bert and traditional models for homophobic content identification on twitter, *arXiv preprint arXiv:2405.09221* (2024).
- [24] F. Barbieri, J. Camacho-Collados, L. Neves, L. E. Anke, Tweeteval: Unified benchmark and comparative evaluation for tweet classification, *CoRR abs/2010.12421* (2020). URL: <https://arxiv.org/abs/2010.12421>. arXiv:2010.12421.
- [25] G. Leoni Santos, V. Gaboardi dos Santos, C. Kearns, G. Sinclair, J. Black, M. Doidge, T. Fletcher, D. Kilvington, K. Liston, P. T. Endo, T. Lynn, Detecting homophobic speech in soccer tweets using large language models and explainable ai, in: *Proc. of the 16th Int. Conf. on Advances in Social Networks Analysis and Mining (ASONAM)*, Lecture Notes in Computer Science, vol. 15211, Springer, 2024, pp. 489–504.
- [26] OpenSearch Project, Opensearch: Open source search and analytics suite, <https://opensearch.org>, 2021. Accessed: 2025-06-02.