VEL@IberLEF 2024:Hope Speech Detection in Spanish Social Media Comments using BERT Pre-trained Model

Kishore Kumar Ponnusamy[1,*], Mani Vegupatti[2], Prasanna Kumar Kumaresan[3], Ruba Priyadharshini[4], Paul Buitelaar[3] and Bharathi Raja Chakravarthi[3]

[1]Digital University Kerala, Kerala, India
[2]University of Galway, Ireland
[3]Data Science Institute, University of Galway, Ireland
[4]Gandhigram Rural Institute-Deemed to be University, India.

Abstract

Our research aims to address the task of detecting hope speech in Spanish text from social media. In times of stress or adversity, hope speech has a powerful impact on fostering equality, diversity, and inclusion (EDI). It has the ability to inspire, offer suggestions, and uplift individuals, making it an important tool in promoting a more inclusive society. Despite its significance, the detection of hope speech in social media data remains relatively unexplored in natural language processing (NLP). To bridge this gap, we present a novel approach leveraging the "bert-base-spanish-wwm-cased" pre-trained model to detect hope speech in Spanish tweets. We preprocess the data and establish a baseline for hope speech detection, emphasising its importance in promoting EDI. Our model achieves 7th rank with a macro F1 score of 0.60, demonstrating promising results in identifying instances of hope speech in both LGTBI and unknown domains. Our research contributes to advancing the understanding of hope speech in Spanish social media, paving the way for future applications in EDI initiatives and beyond.

Keywords

Hope speech detection, Spanish social media texts, Pre-trained transformer, Natural Language Processing, Equality, Diversity, Inclusion (EDI), LGTBI domain

1. Introduction

Over the past few decades, the emergence of social media platforms such as YouTube and Facebook has facilitated the widespread sharing of individuals' thoughts and emotions on the internet, reaching billions of people. According to [1, 2, 3], this opportunity has resulted in both detrimental and beneficial outcomes, as well as the sharing of valuable ideas [4]. The proliferation of social media has given rise to a significant problem that researchers must tackle, as there is an urgent need for automated moderation solutions [5]. The subjective nature of the terminology used on social media can cause significant harm and distress to lesbian, gay, bisexual, transgender, queer, and other associated populations (LGBT). It could have more substantial consequences for these areas [6, 7, 8]. Hope speech can be defined as the cognitive capacity of humans to imagine future events and their potential outcomes while maintaining a flexible perspective [9]. Hope speech recognition is an innovative method to encourage and advance optimistic content that contributes to creating a more tranquil and joyful society [10, 11].

Hope is an important element of life. Hope fuels the growth of humanity and accelerates innovation. Especially for those people in a marginalized society, hope is an important element that can help them move forward and make progress in life [12]. In an era where social media and its influence are prevalent, we must have a mechanism to find out the hope speech and use it as a growth catalyst for promoting human well-being, elevating success in various areas, and reducing negative elements like stress, anxiety, and fear, to mention a few [13, 14].

IberLEF 2024, September 2024, Valladolid, Spain.
*Corresponding author.
kishorep161002@gmail.com (K. K. Ponnusamy); mani.vegupatti@gmail.com (M. Vegupatti); kprasannakumar30@gmail.com (P. K. Kumaresan); rubapriyadharshini.a@gmail.com (R. Priyadharshini); paul.Buitelaar@universityofgalway.ie (P. Buitelaar); bharathi.raja@universityofgalway.ie (B. R. Chakravarthi)
0000-0001-9621-668X (K. K. Ponnusamy); 0009-0005-1322-952X (M. Vegupatti); 0000-0003-2244-246X (P. K. Kumaresan); 0000-0003-2323-1701 (R. Priyadharshini); 0000-0001-7238-9842 (P. Buitelaar); 0000-0002-4575-7934 (B. R. Chakravarthi)

We participated in the shared task named "HOPE": Approaching Hope Speech Detection in Social Media from Two Perspectives, for Equality, Diversity, and Inclusion, and as Expectations [15], which is conducted as part of IberLEF 2024 [1], we secured 7th rank in the rank list for the Spanish language. Our approach got the macro-averaged results of 0.64 precision, 0.62 recall, and 0.60 F1 score.

2. Related Work

In his work, Snyder (2000) [16] asserts that hope is a potent compensating attribute in human psychology, enabling individuals to effectively cope with challenges. Snyder et al. (1991) [17] define hope as a cognitive framework that can be understood as the process of establishing objectives. The feature in question encompasses both individuals' motivation to remain dedicated to a future objective and their capacity to envision a path toward achieving that objective. In contrast, Averill et al. (2012) [18] asserted that hope is an emotion contingent upon an individual's perception of the likelihood of achieving their objective. Various interpretations of hope can be found in the literature.

Distinguishing between text on social media and content published in an optimistic tone can be simultaneously intriguing and challenging. To complete this project, you must record the precise details and language used to depict positive discourse in social media posts. The study conducted by Palakodety et al. (2019) [19] was the first to propose that expressions of optimism may be identified in English YouTube comments. The methodology employed involved the utilization of the Logistic Regression (LR) model, n-grams, mood analysis, and fastText embeddings to analyze the hope speech. Subsequently, Chakravarthi et al. [9] conducted research on identifying hope speech in English, Tamil, and Malayalam. Furthermore, he made the dataset publicly available to facilitate further investigation in this area. A recent study focused on detecting hopeful speech in the Spanish language. This research included creating SpanishHopeEDI, a new Spanish Twitter dataset on the LGBT community, and conducting experiments that serve as a baseline for further research [20]. The study conducted by Balouchzahi et al. (2023) [21] focused on detecting hope speech at a hierarchical level. They categorized hope speech into three types: generalized hope, realistic hope, and unrealistic hope. However, they did not specifically address the categorization of hope speech in several languages.

Shared task organized by Chakravarthi et al. released a dataset for shared task focusing on detecting hope speech in Tamil, English, and Malayalam languages [22]. The shared assignment was carried out as a component of the EACL 2021 workshop on Language Technology for Equality, Diversity, and Inclusion (LT-EDI-2021). As far as we know, this is the first shared task dedicated to the identification of hopeful speech. The second workshop on Language Technology for Equality, Diversity, and Inclusion (LT-EDI-2022) included a shared task on hope speech detection for Tamil, Malayalam, Kannada, English, and Spanish languages [23]. This workshop was organized as a part of ACL 2022. The participants were given training and development datasets that had been annotated, as well as unlabeled test datasets, in all five languages. The IberLEF 2023 shared task [11], HOPE: Multilingual Hope Speech Detection, aimed to determine whether English or Spanish texts contain expressions of hope or not. The competition was facilitated by CodaLab and garnered registrations from 50 teams. In conclusion, there were a total of 12 results submitted and 8 working notes given, which provided detailed descriptions of their respective systems.

The statistical models are still used as baselines for respective tasks by various researchers. Kumaresan et al. [24] explored statistical machine learning (ML) models such as naive Bayes, support vector machine, random forest, decision tree, and logistic regression with feature extraction techniques such as TF-IDF and count vectorizers. They also used Deep Learning (DL) models such as RoBERTa [25], XLM-RoBERTa [26], and Multilingual BERT [27]. These models are transformer architectures that are pre-trained on mono and multiple languages for hope speech detection in Tamil, English, and Tamil-English (code-mixed). With the ML models, the different feature extraction methods were explored by Ponnusamy et al. [28]. They used doc2vec and MLPNet to extract features from the Bulgarian language texts for the hope speech detection task. For the Hope Speech 2022 shared task, Vijayakumar et al. [29] utilized BERT,

[1]https://sites.google.com/view/iberlef-2024/

and their model scored first in Kannada, second in Malayalam, third in Tamil, and sixth in English. We found that "dccuchile/bert-base-spanish-wwm- cased" [30] is pre-trained on a large amount of Spanish text and produced better results on the GLUE benchmark than other BERT-based transformer models. Hence, we chose to use this model and fine-tune it for our task.

3. Task and Dataset Description

Task 1 focuses on "Hope for Equality, Diversity, and Inclusion," highlighting the significance of hopeful language in reducing hostility and providing support to individuals dealing with difficulties like illness, stress, or loneliness. This is especially important for vulnerable groups like the LGBT community and racial minorities. This challenge involves analyzing a Spanish tweet to determine if it contains hopeful rhetoric or not. The classes for each text are:

- HS: hope speech.
- NHS: non-hope speech.

3.1. Dataset Description

The dataset for this task was collected between 2020 and 2023. It is an improved and extended version of the SpanishHopeEDI dataset [20]. Different versions of the dataset were used in the shared tasks on hope speech detection at LT-EDI-2022 [2], as a part of ACL 2022 [23], at LT-EDI-2023 [3], within RANLP 2023 [12] and in the shared task HOPE at IberLEF 2023 [11]. The version of the dataset[4] for IberLEF 2024 [31] consists of training and development sets on LGTBI-related tweets and a test set on tweets related to the LGTBI collective and other EDI topics (unknown domains). A tweet is considered HS if the text: i) explicitly supports the social integration of minorities; ii) is a positive inspiration for the collective; iii) explicitly encourages people who might find themselves in a situation; or iv) unconditionally promotes tolerance. On the contrary, a tweet is marked as NHS if the text: i) expresses negative sentiment towards a community; ii) explicitly seeks violence; or iii) uses gender-based insults. The dataset is composed of 2,000 tweets. The statistics of the dataset are shown in Table 1. Figure 1 1 shows the word cloud for the highly repeated words in the dataset.

Table 1
Data statistics for Spanish in Task 1

Category	Train	Dev	Test
Hope	700	100	200
Non-hope	700	100	200
Total	**1,400**	**200**	**400**

4. Methodology

In this section, we outline our approach to identifying instances of hope speech in Spanish text from social media. We utilized natural language processing techniques, with a primary goal of refining a pre-trained transformer-based model for sequence classification. Since ML models can't able to understand the raw text, we need to give only numerical values as an input to perform any task.

In this task, we used a few traditional machine learning models and for extracting features we used TF-IDF Vectorizer to convert the text into vectors. We experiment with four ML algorithms including Logistic Regression (LR), Multinomial Naive Baye (MNB), Random Forest (RF), and Gradient Boosting (GB) by using scikit-learn library. We utilised “dccuchile/bert-base-spanish-wwm-cased" model [30] is a

[2]https://competitions.codalab.org/competitions/36393
[3]https://codalab.lisn.upsaclay.fr/competitions/11076
[4]https://codalab.lisn.upsaclay.fr/competitions/17714

Figure 1: Word Cloud for Train set

pre-trained transformer-based model specifically designed for processing Spanish text. It's based on the Bidirectional Encoder Representations from Transformers (BERT) architecture [27], which was trained with the Whole Word Masking technique.

While BERT models can be adapted for multiple languages, "dccuchile/bert-base-spanish-wwm-cased" is specifically trained on a large corpus of Spanish text. This makes it well-suited for this task which requires a deep understanding of Spanish language nuances to detect the Hope-Speech in Spanish text. The fine-tuning process required training the BERT model on the training dataset for 5 epochs. We employed the AdamW optimizer alongside a learning rate of 1e-5. To regulate the learning rate schedule, we implemented a linear scheduler with warmup steps. Throughout the training process, the model continuously adjusted its parameters by analyzing the calculated loss, ultimately enhancing its ability to detect hope speech.

5. Results and Discussion

In this, we evaluated the models that we used to detect the hope speech in Spanish text. The performance of these models was assessed using metrics such as Accuracy (ACC), Macro Precision (MP), Macro Recall (MR), Macro F1 (MF1), Weighted Precision (WP), Weighted Recall (WR), and Weighted F1 (WF1) scores. We experimented with four ML models (LR, MNB, RF, and GB) with TfidfVectorizer as a feature extraction technique using the scikit-learn[5] library. The macro-F1 scores for these models for the development set are 0.74 (LR), 0.77 (MNB), 0.64 (RF), and 0.64 (GB).

Table 2
Develpment set results for hope speech in Spanish

Models	ACC	MP	MR	MF1	WP	WR	WF1
BERT	**0.84**	**0.85**	**0.84**	**0.84**	**0.85**	**0.84**	**0.84**
LR	0.75	0.78	0.75	0.74	0.78	0.75	0.74
MNB	0.77	0.78	0.77	0.77	0.78	0.77	0.77
RF	0.65	0.67	0.65	0.64	0.67	0.65	0.64
GB	0.65	0.66	0.65	0.64	0.66	0.65	0.64

Next, we experiment with the BERT pre-trained model "bert-base-spanish-wwm-cased". However, we adapted this model for detecting hope speech in the Spanish language with some fine tuning. Comparing to the results in Table 2, this model outperformed the other models in the development set by achieving

[5] https://scikit-learn.org/

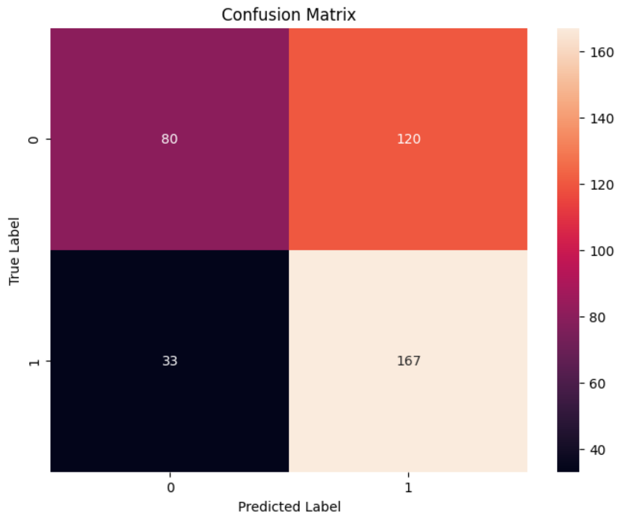

Figure 2: Test Set Confusion matrix for BERT

a macro-F1 score of 0.84. Overall, based on the evaluation and the performance of the BERT pre-trained model "bert-base-spanish-wwm-cased" in the development set, we selected it as the best model for detecting hope-speech Spanish text in social media. The final leaderboard results revealed that the proposed methodology ranked in seventh place in task 1 with a macro F1-score of 0.60. For a better understanding of the model, we visualized the prediction using a confusion matrix, which is shown in figure 2. The visualization provides a clear view of how the model was trained and predicted correctly.

Table 3
Test set results for hope speech in Spanish

Models	ACC	MP	MR	MF1	WP	WR	WF1
BERT	**0.62**	**0.64**	**0.62**	**0.60**	**0.64**	**0.62**	**0.60**
LR	0.57	0.60	0.57	0.54	0.60	0.57	0.54
MNB	0.55	0.55	0.55	0.54	0.55	0.55	0.54
RF	0.57	0.59	0.57	0.54	0.59	0.57	0.54
GB	0.54	0.58	0.54	0.47	0.58	0.54	0.47

6. Conclusion

Our study focused on evaluating the performance of different ML models and the BERT pre-trained model. Our primary objective was to detect hope and non-hope speech in Spanish text. We proposed a methodology that utilises the BERT pre-trained model (bert-base-spanish-wwm-cased) and performed a fine-tuning process to enhance its performance for this purpose.

The proposed model attained a macro-F1 score of 0.60 in task 1, placing 7th, as a result of our endeavours. This demonstrates its proficiency in accurately distinguishing between instances of hope and non-hope discourse in Spanish text. Through the utilization of the BERT pre-trained model and the subsequent optimization of its parameters via fine-tuning, we achieved a notable enhancement in its classification accuracy, thereby making significant improvements to the progress of this particular area of research.

Acknowledgement

Author Bharathi Raja Chakravarthi was supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289_P2(Insight_2). Prasanna Kumar Kumaresan was supported in part by a research grant from the Science Foundation Ireland Centre for Research Training in Artificial Intelligence under Grant No. 18/CRT/6223.

References

[1] B. R. Chakravarthi, V. Muralidaran, R. Priyadharshini, J. P. McCrae, Corpus creation for sentiment analysis in code-mixed Tamil-English text, in: D. Beermann, L. Besacier, S. Sakti, C. Soria (Eds.), Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL), European Language Resources association, Marseille, France, 2020, pp. 202–210. URL: https://aclanthology.org/2020.sltu-1.28.

[2] N. Jindal, P. K. Kumaresan, R. Ponnusamy, S. Thavareesan, S. Rajiakodi, B. R. Chakravarthi, Mistra: Misogyny detection through text–image fusion and representation analysis, Natural Language Processing Journal 7 (2024) 100073. URL: https://www.sciencedirect.com/science/article/pii/S2949719124000219. doi:https://doi.org/10.1016/j.nlp.2024.100073.

[3] V. S. Raj, C. N. Subalalitha, L. Sambath, F. Glavin, B. R. Chakravarthi, Conbert-rl: A policy-driven deep reinforcement learning based approach for detecting homophobia and transphobia in low-resource languages, Natural Language Processing Journal 6 (2024) 100040. URL: https://www.sciencedirect.com/science/article/pii/S2949719123000377. doi:https://doi.org/10.1016/j.nlp.2023.100040.

[4] 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 Journal 5 (2023) 100041. URL: https://www.sciencedirect.com/science/article/pii/S2949719123000389. doi:https://doi.org/10.1016/j.nlp.2023.100041.

[5] B. R. Chakravarthi, A. Hande, R. Ponnusamy, P. K. Kumaresan, R. Priyadharshini, How can we detect homophobia and transphobia? experiments in a multilingual code-mixed setting for social media governance, International Journal of Information Management Data Insights 2 (2022) 100119. URL: https://www.sciencedirect.com/science/article/pii/S2667096822000623. doi:https://doi.org/10.1016/j.jjimei.2022.100119.

[6] B. R. Chakravarthi, R. Priyadharshini, S. Banerjee, M. B. Jagadeeshan, P. K. Kumaresan, R. Ponnusamy, S. Benhur, J. P. McCrae, Detecting abusive comments at a fine-grained level in a low-resource language, Natural Language Processing Journal 3 (2023) 100006. URL: https://www.sciencedirect.com/science/article/pii/S2949719123000031. doi:https://doi.org/10.1016/j.nlp.2023.100006.

[7] M. Subramanian, R. Ponnusamy, S. Benhur, K. Shanmugavadivel, A. Ganesan, D. Ravi, G. K. Shanmugasundaram, R. Priyadharshini, B. R. Chakravarthi, Offensive language detection in tamil youtube comments by adapters and cross-domain knowledge transfer, Computer Speech Language 76 (2022) 101404. URL: https://www.sciencedirect.com/science/article/pii/S0885230822000407. doi:https://doi.org/10.1016/j.csl.2022.101404.

[8] B. R. Chakravarthi, M. B. Jagadeeshan, V. Palanikumar, R. Priyadharshini, Offensive language identification in dravidian languages using mpnet and cnn, International Journal of Information Management Data Insights 3 (2023) 100151. URL: https://www.sciencedirect.com/science/article/pii/S2667096822000945. doi:https://doi.org/10.1016/j.jjimei.2022.100151.

[9] B. R. Chakravarthi, HopeEDI: A multilingual hope speech detection dataset for equality, diversity, and inclusion, in: M. Nissim, V. Patti, B. Plank, E. Durmus (Eds.), Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media, Association for Computational Linguistics, Barcelona, Spain (Online), 2020, pp. 41–53. URL: https://aclanthology.org/2020.peoples-1.5.

[10] B. R. Chakravarthi, Hope speech detection in youtube comments, Social Network Analysis and Mining 12 (2022) 75.

[11] S. M. Jiménez-Zafra, M. Á. Garcia-Cumbreras, D. García-Baena, J. A. Garcia-Díaz, B. R. Chakravarthi,

R. Valencia-García, L. A. Ureña-López, Overview of hope at iberlef 2023: Multilingual hope speech detection, Procesamiento del Lenguaje Natural 71 (2023) 371–381.
[12] P. K. Kumaresan, B. R. Chakravarthi, S. Cn, M. Á. García-Cumbreras, S. M. Jiménez Zafra, J. A. García-Díaz, R. Valencia-García, M. Hardalov, I. Koychev, P. Nakov, D. García-Baena, K. K. Ponnusamy, Overview of the shared task on hope speech detection for equality, diversity, and inclusion, in: B. R. Chakravarthi, B. Bharathi, J. Griffith, K. Bali, P. Buitelaar (Eds.), Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion, INCOMA Ltd., Shoumen, Bulgaria, Varna, Bulgaria, 2023, pp. 47–53. URL: https://aclanthology.org/2023.ltedi-1.7.
[13] H. RamakrishnaIyer LekshmiAmmal, M. Ravikiran, G. Nisha, N. Balamuralidhar, A. Madhusoodanan, A. Kumar Madasamy, B. R. Chakravarthi, Overlapping word removal is all you need: revisiting data imbalance in hope speech detection, Journal of Experimental & Theoretical Artificial Intelligence (2023) 1–23.
[14] B. R. Chakravarthi, Multilingual hope speech detection in english and dravidian languages, International Journal of Data Science and Analytics 14 (2022) 389–406.
[15] D. García-Baena, F. Balouchzahi, S. Butt, M. Á. García-Cumbreras, A. Lambebo Tonja, J. A. García-Díaz, S. Bozkurt, B. R. Chakravarthi, H. G. Ceballos, V.-G. Rafael, G. Sidorov, L. A. Ureña-López, A. Gelbukh, S. M. Jiménez-Zafra, Overview of HOPE at IberLEF 2024: Approaching Hope Speech Detection in Social Media from Two Perspectives, for Equality, Diversity and Inclusion and as Expectations, Procesamiento del Lenguaje Natural 73 (2024).
[16] C. R. Snyder, Handbook of hope: Theory, measures, and applications, Academic press, 2000.
[17] C. R. Snyder, C. Harris, J. R. Anderson, S. A. Holleran, L. M. Irving, S. T. Sigmon, L. Yoshinobu, J. Gibb, C. Langelle, P. Harney, The will and the ways: development and validation of an individual-differences measure of hope., Journal of personality and social psychology 60 (1991) 570.
[18] J. R. Averill, G. Catlin, K. K. Chon, Rules of hope, Springer Science & Business Media, 2012.
[19] S. Palakodety, A. R. KhudaBukhsh, J. G. Carbonell, Hope speech detection: A computational analysis of the voice of peace, arXiv preprint arXiv:1909.12940 (2019).
[20] D. García-Baena, M. Á. García-Cumbreras, S. M. Jiménez-Zafra, J. A. García-Díaz, R. Valencia-García, Hope speech detection in spanish: The lgbt case, Language Resources and Evaluation 57 (2023) 1487–1514.
[21] F. Balouchzahi, G. Sidorov, A. Gelbukh, Polyhope: Two-level hope speech detection from tweets, Expert Systems with Applications 225 (2023) 120078.
[22] B. R. Chakravarthi, V. Muralidaran, Findings of the shared task on hope speech detection for equality, diversity, and inclusion, in: B. R. Chakravarthi, J. P. McCrae, M. Zarrouk, K. Bali, P. Buitelaar (Eds.), Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion, Association for Computational Linguistics, Kyiv, 2021, pp. 61–72. URL: https://aclanthology.org/2021.ltedi-1.8.
[23] B. R. Chakravarthi, V. Muralidaran, R. Priyadharshini, S. Cn, J. McCrae, M. Á. García, S. M. Jiménez-Zafra, R. Valencia-García, P. Kumaresan, R. Ponnusamy, D. García-Baena, J. García-Díaz, Overview of the shared task on hope speech detection for equality, diversity, and inclusion, in: B. R. Chakravarthi, B. Bharathi, J. P. McCrae, M. Zarrouk, K. Bali, P. Buitelaar (Eds.), Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion, Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 378–388. URL: https://aclanthology.org/2022.ltedi-1.58. doi:10.18653/v1/2022.ltedi-1.58.
[24] P. K. Kumaresan, R. Ponnusamy, E. Sherly, S. Sivanesan, B. R. Chakravarthi, Transformer based hope speech comment classification in code-mixed text, in: International Conference on Speech and Language Technologies for Low-resource Languages, Springer, 2022, pp. 120–137.
[25] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, arXiv preprint arXiv:1907.11692 (2019).
[26] 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.
[27] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers

for language understanding, CoRR abs/1810.04805 (2018). URL: http://arxiv.org/abs/1810.04805. arXiv:1810.04805.
[28] R. Ponnusamy, M. S, S. Thavareesan, R. Priyadharshini, B. R. Chakravarthi, VEL@LT-EDI-2023: Automatic detection of hope speech in Bulgarian language using embedding techniques, in: B. R. Chakravarthi, B. Bharathi, J. Griffith, K. Bali, P. Buitelaar (Eds.), Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion, INCOMA Ltd., Shoumen, Bulgaria, Varna, Bulgaria, 2023, pp. 179–184. URL: https://aclanthology.org/2023.ltedi-1.27.
[29] P. Vijayakumar, P. S, A. P, A. S, R. Sivanaiah, S. M. Rajendram, M. T T, SSN_ARMM@ LT-EDI -ACL2022: Hope speech detection for equality, diversity, and inclusion using ALBERT model, in: B. R. Chakravarthi, B. Bharathi, J. P. McCrae, M. Zarrouk, K. Bali, P. Buitelaar (Eds.), Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion, Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 172–176. URL: https://aclanthology.org/2022.ltedi-1.22. doi:10.18653/v1/2022.ltedi-1.22.
[30] 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.
[31] L. Chiruzzo, S. M. Jiménez-Zafra, F. Rangel, Overview of IberLEF 2024: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2024), co-located with the 40th Conference of the Spanish Society for Natural Language Processing (SEPLN 2024), CEUR-WS.org, 2024.