

# Multilingual Hope Speech Detection using ModernBERT

Michael Ibrahim

Computer Engineering Department, Cairo University, 1 Gamaa Street, 12613, Giza, Egypt

## Abstract

Detecting hope speech in social media texts—a task complicated by sarcasm, multilingual nuances, and contextual ambiguity—is critical for applications in mental health monitoring and inclusive content moderation. This paper presents a fine-tuned ModernBERT model, a state-of-the-art encoder model with extended context windows and hardware optimizations, used to address the PolyHope Shared Task at IberLEF 2025. We fine-tuned ModernBERT for four subtasks: binary and multiclass hope speech detection in English and Spanish. Our approach achieved 1st place in binary English detection with a macro F1-score of **0.872**, outperforming 30 competing teams. For multiclass English (5th place, F1: **0.742**), binary Spanish (6th place, F1: **0.790**), and multiclass Spanish (6th place, F1: **0.655**), the model demonstrated robust cross-lingual transfer capabilities despite limited sarcasm annotations and cultural variations. Results highlight ModernBERT’s efficiency (57-minute training on an NVIDIA T4 GPU) and its superiority in binary classification, while underscoring challenges in multilingual sarcasm detection. This work advances emotion-aware NLP systems by providing a scalable framework for nuanced, real-time hope speech analysis, with implications for equity-focused content moderation and mental health support platforms.

## Keywords

Hope speech detection, ModernBERT, Multilingual NLP, Transformer models

## 1. Introduction

The rapid evolution of natural language processing (NLP) has unlocked unprecedented capabilities in understanding human emotions, but detecting nuanced emotional states like hope remains a formidable challenge. Hope, characterized by its duality of optimism and expectation, is further complicated by contextual factors such as sarcasm, cultural nuances, and multilingual expressions [1]. This paper addresses these challenges through the fine-tuning of ModernBERT, a state-of-the-art encoder model, as a part of the PolyHope Shared Task at IberLEF 2025 [2, 3]. This task focuses on classifying hope speech in English and Spanish social media texts, with subtasks ranging from binary detection to multiclass categorization that includes sarcasm as a distinct class. The developed approach leverages ModernBERT’s architectural advancements [4], such as extended context windows, hardware optimizations, and diverse pretraining data, to achieve robust performance in detecting nuanced hope expressions while addressing computational efficiency, a critical factor in real-world deployments.

Hope is a multifaceted emotion deeply intertwined with human decision-making, mental health, and social interactions<sup>13</sup>. Unlike basic emotions like joy or anger, hope often coexists with contradictory signals, such as frustration or sarcasm, making it inherently ambiguous for computational systems. For instance, the statement “I hope this situation gets better” might reflect genuine optimism or veiled resignation depending on context. Traditional sentiment analysis frameworks, which categorize text into broad positive, negative, or neutral labels, fail to capture these subtleties. This limitation is exacerbated in multilingual settings, where linguistic and cultural differences further obscure emotional intent [1].

The PolyHope dataset (V2), used in this shared task, addresses these gaps by providing over 30,000 annotated tweets in English and Spanish, distinguishing between four hope subtypes: Generalized Hope, Realistic Hope, Unrealistic Hope, and Sarcasm. This granularity reflects the need for models to discern not only hopeful intent but also its pragmatic and contextual underpinnings. The main challenges of sarcasm detection in hope speech, where surface-level positivity often masks underlying negativity—a problem compounded by the lack of large, annotated datasets.

---

*IberLEF 2025, September 2025, Zaragoza, Spain*

✉ michael.nawar@eng.cu.edu.eg (M. Ibrahim)

id 0000-0003-2340-8917 (M. Ibrahim)



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

While large language models (LLMs) like GPT-4 [5] and Llama 3 [6] have demonstrated remarkable generative capabilities, their computational cost and latency render them impractical for high-throughput classification tasks. Encoder-only models like BERT [7], by contrast, excel in efficiency and accuracy for discriminative tasks, making them indispensable in production pipelines for content moderation, recommendation systems, and retrieval-augmented generation (RAG) [8]. ModernBERT, introduced in late 2024, represents a paradigm shift in encoder design. With architectural innovations such as rotary positional embeddings (RoPE), flash attention optimization, and an 8192-token context window, ModernBERT achieves state-of-the-art performance on classification benchmarks while reducing inference latency by 2–4× compared to BERT [4].

ModernBERT’s pretraining on 2 trillion tokens of diverse data—including web documents, code, and scientific articles—enhances its ability to generalize across domains, a critical advantage for social media text analysis. For example, its exposure to code enables better parsing of informal language and syntactic variations common in platforms like Twitter. These features position ModernBERT as an ideal candidate for the PolyHope task, where detecting sarcasm and culturally specific hope expressions demands both contextual depth and computational agility.

The PolyHope task’s bilingual focus (English and Spanish) introduces unique challenges. Cross-linguistic variations in expressing hope, such as the use of subjunctive moods in Spanish or idiomatic phrases in English, require the models to adapt to syntactic and semantic divergences. Furthermore, sarcasm detection hinges on pragmatic cues that are often language-specific. For instance, Spanish sarcasm frequently relies on irony markers like “claro que sí” (of course), which lack direct equivalents in English.

Existing approaches to hope speech detection, such as the HopeEDI dataset for equality and diversity contexts, have primarily focused on binary classification, neglecting sarcasm and hope subtypes [9]. The IberLEF 2023 shared task made initial strides in Spanish hope detection but lacked the granularity of PolyHope V2 [10]. Recent studies have highlighted the importance of developing multilingual models capable of cross-lingual transfer learning; however, progress has been constrained by a shortage of sarcasm annotations and the complexities of code-mixed data [11].

The developed model builds on the hypothesis that ModernBERT’s architectural advancements address key limitations of earlier models in the domain of multilingual sentiment and figurative language understanding. One major improvement is its extended context awareness, achieved through an 8192-token window, which allows the model to capture broader discourse patterns—an essential capability for detecting nuanced phenomena such as sarcasm and unrealistic hope in long-form social media threads [12]. Additionally, ModernBERT benefits from multilingual pretraining on a diverse corpus that includes Spanish text, thereby mitigating domain shift issues commonly faced in zero-shot cross-lingual transfer scenarios [13]. Finally, the incorporation of flash attention and other hardware-aware optimizations significantly enhances training and inference efficiency, making it feasible to scale experiments to large benchmark datasets such as PolyHope V2 [14].

The remainder of this paper is structured as follows: Section 2 reviews related work on hope speech detection and encoder architectures. Section 3 details our methodology, including dataset preprocessing and hyperparameter configurations. Section 4 presents our results on the PolyHope. Section 5 concludes the work with future directions.

## 2. Related Work

The detection of hope speech and its nuanced subtypes—particularly in multilingual contexts with sarcasm—builds upon advancements in emotion recognition, transformer-based architectures, and cross-linguistic NLP. This section synthesizes prior research into three key categories: (1) hope speech detection and sarcasm analysis, (2) multilingual and cross-lingual emotion recognition, and (3) encoder models in emotion classification. By contextualizing our work within these domains, we highlight the gaps our approach addresses and the innovations enabled by ModernBERT.

## 2.1. Hope Speech Detection and Sarcasm Analysis

Hope, as a complex emotion, has historically been underrepresented in NLP research compared to simpler sentiment categories like positivity or negativity. Early work introduced the HopeEDI dataset [9], focusing on hope speech in equality and diversity contexts, but limited its scope to binary classification (hopeful vs. non-hopeful) in English. Subsequent efforts, such as the shared task IberLEF 2023 [10], expanded to Spanish hope detection but retained a binary framework, failing to address the subtypes of sarcasm or granular hope.

The PolyHope V2 dataset [14] marked a critical shift by proposing a two-level annotation framework: distinguishing Generalized Hope, Realistic Hope, and Unrealistic Hope, while also incorporating sarcasm as a distinct class. This work laid the groundwork for the PolyHope Shared Task at IberLEF 2025 [2] which focuses on detecting optimism, expectation, and sarcasm in English and Spanish social media texts. The task’s multi-class structure addresses prior limitations in dataset granularity, enabling models to capture pragmatic ambiguities where sarcasm masks genuine hope or vice versa.

Sarcasm detection in emotional contexts presents unique challenges, as sarcasm often masks underlying sentiments through irony or cultural context. Sarcasm in hope speech frequently correlates with unrealistic expectations or veiled frustration [15]. Building on this, a transformer-based framework was proposed for multiclass hope speech detection, explicitly modeling sarcasm as a pragmatic subtype in both Spanish and English [16]. The results underscored the necessity of large, annotated datasets like PolyHope V2, which provides over 30,000 tweets with fine-grained labels.

Prior shared tasks, such as hope at IberLEF 2024 [17], explored hope speech detection for equality, diversity, and inclusion (EDI) contexts, framing hope as both a social catalyst and an expression of expectations. While this work advanced the field, its binary classification approach and monolingual focus (Spanish) limited its applicability to multilingual, sarcasm-aware systems.

## 2.2. Multilingual and Cross-Lingual Emotion Recognition

Multilingual emotion detection requires models to navigate complex syntactic, semantic, and cultural differences. MIND-HOPE [18], a multilingual framework designed to identify nuanced dimensions of hope across six languages, underscores the importance of cultural context in distinguishing different types of hope. It demonstrated that models pretrained primarily on English often perform poorly in languages like Spanish, where grammatical features such as the subjunctive mood are crucial for expressing hypothetical or aspirational statements. Additionally, research on hope speech within Spanish-speaking LGBTQ+ communities has shown that cultural and linguistic markers—such as irony and the use of reclaimed slurs—play a significant role in shaping emotional expression [19].

The PolyHope task builds on these insights by introducing sarcasm detection as a language-specific challenge, emphasizing the nuanced ways in which hope and related emotions are expressed across different linguistic contexts. Additional research on hope in Urdu has shown that expressions of hope and hopelessness often rely on metaphorical language that does not appear in Indo-European languages, further underscoring the complexity of cross-lingual emotion detection [20]. Together, these developments highlight the importance of designing models that can adapt to both syntactic and cultural divergences—a challenge that ModernBERT’s multilingual pretraining begins to address.

## 2.3. Encoder Models in Emotion Classification

Encoder-only models like BERT have dominated discriminative NLP tasks due to their efficiency and accuracy. BERT [7] was proven efficient in sentiment analysis, but its fixed 512-token context window hindered performance on longer social media threads. Subsequent variants, such as RoBERTa [21] and DeBERTaV3 [22], improved robustness through dynamic masking and disentangled attention but incurred higher computational costs.

The work of [16] evaluated transformer architectures on hope and regret speech detection, revealing that models like BERT and RoBERTa struggle with class imbalance and sarcasm. Their work highlighted the need for architectures with extended context windows to capture discourse-level cues, a

challenge addressed by ModernBERT's 8192-token capacity. ModernBERT, introduced in late 2024, integrates rotary positional embeddings (RoPE) and flash attention optimization, achieving state-of-the-art performance on emotion classification while reducing inference latency by 2–4× compared to BERT.

While prior research has laid important groundwork in multilingual emotion detection, three key limitations remain. First, existing datasets such as those used in earlier tasks have typically treated hope as a single, undifferentiated category, overlooking critical subtypes like Unrealistic Hope and their nuanced interaction with elements like sarcasm. Second, although state-of-the-art models such as DeBERTaV3 deliver strong performance, they are often computationally intensive and unsuitable for real-time analysis of fast-moving social media streams due to high latency. Third, many multilingual models struggle with cross-lingual generalization, failing to capture language-specific sarcasm cues and thereby performing poorly when transferring between languages like English and Spanish. Our methodology addresses these challenges by leveraging ModernBERT's extended context window, hardware-efficient architecture, and multilingual pretraining. Through fine-tuning, the developed ModernBERT model surpasses baseline models in both accuracy and efficiency, establishing a new standard for nuanced, real-world emotion detection.

### 3. Methodology

This study applied a unified methodological framework to four subtasks of the PolyHope Shared Task: binary hope detection (English), multiclass hope detection (English), binary hope detection (Spanish), and multiclass hope detection (Spanish). The same preprocessing pipeline, model architecture, training protocol, and evaluation metrics were employed across all tasks, ensuring consistency and comparability. Language-specific adjustments were limited to dataset selection (English/Spanish splits) and label mapping (binary: Hope/Not Hope; multiclass: Sarcasm, Generalized Hope, Realistic Hope, Unrealistic Hope). For Spanish tasks, the model leveraged ModernBERT's multilingual pretraining corpus, which included Spanish text, avoiding the need for language-specific tokenizers or architectural modifications.

The English dataset comprised 7,135 annotated English tweets divided into training (5,233 samples) and development (1,902 samples) sets, labeled across five categories: Not Hope, Sarcasm, Generalized Hope, Realistic Hope, and Unrealistic Hope. Preprocessing involved converting categorical labels into numerical identifiers to facilitate model training. Class imbalance was observed, particularly for Unrealistic Hope and Sarcasm categories. While explicit oversampling techniques were not applied, the training process prioritized macro F1-score optimization to mitigate bias toward majority classes. The dataset was structured into Hugging Face DatasetDict objects to streamline integration with transformer-based pipelines.

The base model, ModernBERT-base, was selected for its enhanced capabilities, including an extended context window (8,192 tokens) and hardware-optimized attention mechanisms. Key architectural features included rotary positional embeddings (RoPE), which improved handling of variable-length sequences, and flash attention, which accelerated inference speed by 2–4× compared to standard BERT. Tokenization preserved social media-specific markers (e.g., #USER#) to maintain discourse context.

Fine-tuning was performed using Hugging Face's Trainer API with a set of hyperparameters optimized for stability and efficiency. A training batch size of 32 and an evaluation batch size of 16 were selected to balance GPU memory constraints with gradient stability. The learning rate was set to 5e-5 to enable timely convergence while minimizing the risk of overshooting during gradient updates. Mixed-precision training using bfloat16 was employed to take advantage of hardware acceleration on an NVIDIA T4 GPU with 16GB of VRAM. For optimization, the fused AdamW optimizer was used to enhance memory efficiency, and gradient clipping with a maximum norm of 1.0 was applied to prevent exploding gradients. Training was conducted over 8 epochs, with validation carried out at the end of each epoch to monitor for performance drift.

Model performance was evaluated using the macro F1-score, chosen for its robustness to class imbalance across the five emotion categories. Validation loss was tracked in parallel to monitor for signs of overfitting. Iterative tuning was guided by performance on the development set, and the

best-performing model checkpoint was retained based on peak F1 performance.

Predictions on the test set were generated through an iterative inference pipeline designed to handle variable-length tweets. Each text input was tokenized individually, passed through the model, and classified into one of the predefined categories. Predicted numerical labels were subsequently mapped to their semantic equivalents (e.g., 2  $\rightarrow$  Generalized Hope), producing outputs suitable for human interpretation and qualitative analysis.

The full training cycle completed in 57 minutes, achieving a processing throughput of 12.23 samples per second. ModernBERT’s use of flash attention and fused operations significantly reduced memory overhead, enabling efficient use of the NVIDIA A100 GPU. This was accomplished despite the model’s substantial architecture, which includes 22 layers and 768-dimensional hidden states.

Two key limitations were identified during evaluation. First, class imbalance remained a challenge; although macro F1 was optimized, underrepresented classes such as Sarcasm showed slightly lower recall, indicating that future work may benefit from the use of weighted loss functions or data augmentation strategies. Second, the model’s extended 8k-token context window was not fully leveraged when processing short-form content like tweets, suggesting a need to explore dynamic truncation or chunking strategies better suited to long-form social media threads.

## 4. Results

Table 1 summarizes the final test-set results for the four PolyHope subtasks. ModernBERT achieves very high F1-scores on the binary hope-detection task in both languages (English F1 = **0.872**, Spanish = **0.790**). This indicates consistently strong accuracy in distinguishing “hope” vs. “not-hope.” These results are in line with prior fine-tuned transformer models, which have reported macro-F1 in the mid-80’s on English hope-speech binary classification. The results also show small divergence between weighted and macro scores ( $< 0.3\%$ ) suggests the model treats both classes (Hope and Not-Hope) nearly equally. In other words, ModernBERT’s performance is balanced across languages and labels, confirming that the broad binary prediction task is handled robustly.

Task	Language	Accuracy	Precision	Recall	F1-score
Binary (test)	English	0.872	0.872	0.872	0.872
	Spanish	0.790	0.790	0.790	0.790
Multiclass (test)	English	0.782	0.748	0.782	0.781
	Spanish	0.705	0.697	0.705	0.700

**Table 1**

Test set performance of ModernBERT on the PolyHope tasks.

In the more challenging multiclass tasks, as expected, performance drops when moving to the finer-grained 5-way task. ModernBERT attains  $F1 = 0.782$  on English and 0.700 on Spanish for the multiclass (Generalized / Realistic / Unrealistic / Sarcastic / Not-Hope) task. This roughly 10~12% absolute decrease relative to the binary case is consistent with the added difficulty of distinguishing multiple nuanced categories. Importantly, these multiclass scores remain solid (well above chance) and the drop is moderate. This suggests the model’s understanding is still good overall, even though the classification boundary becomes more complex.

Overall, the results paint a balanced picture. ModernBERT clearly excels at the simpler binary task, confirming its capacity for broad hope detection. The drop in multiclass performance, while expected, highlights areas for further work. Specifically, the modest decline (10~12%) signals that the model still captures much of the fine-grained structure, but struggles with subtler distinctions. This underscores the need for future improvements in multilingual fine-grained emotion classification, for example, more training data or specialized modeling for sarcastic and highly optimistic utterances. Going forward, augmenting context understanding or incorporating pragmatic cues may help ModernBERT better disambiguate Sarcasm and Unrealistic Hope.

## 5. Conclusion and Future Work

This study demonstrated the effectiveness of ModernBERT—a model optimized for extended context and hardware efficiency—in detecting hope speech across different languages and task complexities. ModernBERT achieved first place in the binary English detection task with a macro F1 score of 0.872, alongside competitive results in multiclass and Spanish tasks. These outcomes validated the model’s strengths in binary classification and cross-lingual transfer. Key innovations that contributed to these results included the use of rotary positional embeddings to capture nuanced contextual information and flash attention, which facilitated rapid inference, making the model suitable for real-time applications. However, performance gaps in sarcasm detection, such as a multiclass Spanish F1 score of 0.654, and issues related to class imbalance highlighted the challenges of modeling pragmatic and cultural subtleties inherent in emotion detection.

In order to address the identified limitations, several future directions are proposed. First, sarcasm-aware architectures could be developed by integrating linguistic markers (e.g., irony indicators) or multimodal cues, such as emojis and images, to enhance sarcasm detection, particularly in low-resource languages. Additionally, to fully exploit ModernBERT’s 8k-token window, dynamic context utilization strategies should be explored, adapting the model for platforms with long-form content like Reddit threads. For handling class imbalance, hybrid approaches involving focal loss, synthetic data generation (e.g., using large language models to augment sarcasm samples), and adversarial training could be tested. Another avenue for future work is the multilingual expansion of ModernBERT, particularly evaluating its performance on under-resourced languages such as Urdu and Arabic using frameworks like UrduHope, with attention to dialectal and script diversity. Efficiency optimization can also be pursued through model distillation or quantization techniques to deploy lightweight variants for real-time content moderation on edge devices. Finally, to better capture hope expressions in diverse sociolinguistic contexts, cultural nuance modeling could be enhanced by incorporating region-specific lexicons or leveraging collaborative filtering techniques.

By addressing these challenges, future work can advance emotion-aware NLP systems, promoting global and equitable applications in areas such as mental health support and creating inclusive online spaces.

## Declaration on Generative AI

Generative AI (LLMs) assisted in improving textual coherence and readability. All ideas, experiments, and conclusions originated from the authors. No AI contribution influenced the study’s findings or evidence.

## References

- [1] S. Butt, F. Balouchzahi, A. I. Amjad, M. Amjad, H. G. Ceballos, S. M. Jimenez-Zafra, Optimism, expectation, or sarcasm? multi-class hope speech detection in spanish and english, arXiv preprint arXiv:2504.17974 (2025).
- [2] S. Butt, F. Balouchzahi, M. Amjad, S. M. Jimenez-Zafra, H. G. Ceballos, G. Sidorov, Overview of polyhope at iberlef 2025: Optimism, expectation or sarcasm?, in: Procesamiento del Lenguaje Natural., 2025.
- [3] 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.
- [4] B. Warner, A. Chaffin, B. Clavié, O. Weller, O. Hallström, S. Taghadouini, A. Gallagher, R. Biswas, F. Ladhak, T. Aarsen, et al., Smarter, better, faster, longer: A modern bidirectional encoder for fast,



memory efficient, and long context finetuning and inference, arXiv preprint arXiv:2412.13663 (2024).

- [5] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., Gpt-4 technical report, arXiv preprint arXiv:2303.08774 (2023).
- [6] A. Grattafiori, A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Vaughan, et al., The llama 3 herd of models, arXiv preprint arXiv:2407.21783 (2024).
- [7] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, 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), 2019, pp. 4171–4186.
- [8] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., Retrieval-augmented generation for knowledge-intensive nlp tasks, *Advances in neural information processing systems* 33 (2020) 9459–9474.
- [9] B. R. Chakravarthi, Hopeedi: A multilingual hope speech detection dataset for equality, diversity, and inclusion, in: Proceedings of the Third Workshop on Computational Modeling of People’s Opinions, Personality, and Emotion’s in Social Media, 2020, pp. 41–53.
- [10] S. M. Jiménez-Zafra, M. Á. García-Cumbreras, D. García-Baena, J. A. García-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.
- [11] B. R. Chakravarthi, V. Muralidaran, R. Priyadharshini, S. Cn, J. P. McCrae, M. Á. García, S. M. Jiménez-Zafra, R. Valencia-García, P. Kumaresan, R. Ponnusamy, et al., Overview of the shared task on hope speech detection for equality, diversity, and inclusion, in: Proceedings of the second workshop on language technology for equality, diversity and inclusion, 2022, pp. 378–388.
- [12] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, Exploring the limits of transfer learning with a unified text-to-text transformer, *Journal of machine learning research* 21 (2020) 1–67.
- [13] A. Conneau, A. Baevski, R. Collobert, A. Mohamed, M. Auli, Unsupervised cross-lingual representation learning for speech recognition, arXiv preprint arXiv:2006.13979 (2020).
- [14] F. Balouchzahi, G. Sidorov, A. Gelbukh, Polyhope: Two-level hope speech detection from tweets, *Expert Systems with Applications* 225 (2023) 120078.
- [15] F. Balouchzahi, S. Butt, G. Sidorov, A. Gelbukh, Cic@ It-edi-acl2022: Are transformers the only hope? hope speech detection for spanish and english comments, in: Proceedings of the second workshop on language technology for equality, diversity and inclusion, 2022, pp. 206–211.
- [16] G. Sidorov, F. Balouchzahi, S. Butt, A. Gelbukh, Regret and hope on transformers: An analysis of transformers on regret and hope speech detection datasets, *Applied Sciences* 13 (2023) 3983.
- [17] D. García-Baena, F. Balouchzahi, S. Butt, M. Á. García-Cumbreras, A. L. Tonja, J. A. García-Díaz, S. Bozkurt, B. R. Chakravarthi, H. G. Ceballos, R. Valencia-García, et al., 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) 407–419.
- [18] G. Sidorov, F. Balouchzahi, L. Ramos, H. Gómez-Adorno, A. Gelbukh, Mind-hope: Multilingual identification of nuanced dimensions of hope (2024).
- [19] 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.
- [20] F. Balouchzahi, S. Butt, M. Amjad, G. Sidorov, A. Gelbukh, Urduhope: Analysis of hope and hopelessness in urdu texts, *Knowledge-Based Systems* 308 (2025) 112746.
- [21] 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).
- [22] P. He, J. Gao, W. Chen, Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing, arXiv preprint arXiv:2111.09543 (2021).