

HOPE: A Multilingual Approach to Identifying Positive Communication in Social Media

Fida Ullah¹, Muhammad Tayyab Zamir¹, Muhammad Ahmad¹, Grigori Sidorov¹ and Alexander Gelbukh¹

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

Abstract

The process of identifying hope speech involves classifying sentences into those that convey hopeful messages and those that do not, based on a dataset of text data. Hopeful expression encompasses positive, supportive, inclusive, and reassuring communications aimed at inspiring optimism in individuals. Unlike the focus on recognizing and reducing negative language usage, detecting hope speech aims to discover and enhance positive modes of communication in online interactions. In our paper, we detail our participation in the HOPE: Multilingual Hope Speech Detection shared task at IberLEF 2024. This task includes two sub-tasks: identifying hope speech in Spanish and English tweets sourced from social media content. Our approach with BERT multilingual employs a word-based tokenization strategy for training which yielded an F1 Score of 0.71 for Spanish and 0.74 for English language.

Keywords

Hope, Multilingual, BERT, Tokenization, Natural Language Processing

1. Introduction

In today's digitally interconnected world, social media platforms have become vital arenas for communication, where individuals express a wide array of sentiments [1, 2] and opinions. Among these expressions, hope speech [3, 4] stands out as a beacon of optimism, fostering positivity and resilience within communities facing adversity. Understanding and detecting hope speech in online discourse is paramount, as it not only reflects the collective mindset of societies but also offers insights into the dynamics of social interaction and psychological [5, 6] well-being. The rise of computational linguistics and natural language processing (NLP) methods has provided opportunities for automated examination of textual information, empowering researchers to explore the nuances of online human communication. They engaged in diverse tasks with the dataset, such as scrutinizing fake news [7] pinpointing hate speech [8, 9], recognizing language structures, performing sentiment analysis, and investigating expressions of optimism. These endeavors encompassed a thorough exploration of the data to uncover insights regarding misinformation, language identification [10, 11] linguistic variations, emotional nuances, and hopeful expressions within the dataset. Within this context, the identification and classification of hope speech present a significant challenge, given its nuanced and context-dependent nature. However, the potential benefits of such research are manifold, ranging from enhancing mental health interventions to fostering a more optimistic and supportive online environment. Social networking platforms like Facebook, Twitter, Instagram, and YouTube have attracted a large user base and provide a platform for content sharing and opinion expression. They also play an important role in offering support to marginalized communities [5]. With the challenges posed by the pandemic causing disruptions across various essential services, people are turning to online forums to fulfill their informational, emotional, and social needs. These platforms not only facilitate networking but also contribute to individuals' sense of belonging within their communities [12, 13, 14]. Despite their positive impact on mental health and well-being, social media platforms also host a significant amount of negative or harmful content

IberLEF 2024, September 2024, Valladolid, Spain

✉ Fidaullahmohmand@gmail.com (F. Ullah); tayyab.awan8001@gmail.com (M. T. Zamir); mahmad.riaz102@gmail.com (M. Ahmad); sidorov@cic.ipn.mx (G. Sidorov); gelbukh@cic.ipn.mx (A. Gelbukh)

🆔 0000-0003-3901-3522 (G. Sidorov); 0000-0001-7845-9039 (A. Gelbukh)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

due to inadequate moderation mechanisms [8]. Efforts have been made to address this issue through techniques such as hate speech detection [15] and identifying offensive languages [16].

2. Literature Review

In today's interconnected world, the detection of hope speech in multilingual contexts has become increasingly important [17]. Languages are a fundamental part of our identity and culture, and understanding hope speech across different languages can provide valuable insights into the human experience [16]. However, detecting hope speech in different languages presents unique challenges [11]. The ability to identify hope speech has significant implications in social media platforms, where understanding and promoting positive and uplifting content can contribute to a healthier online environment [5, 18]. Additionally, the development of datasets for hope speech in multiple languages is essential for the creation of effective detection models. By evaluating the effectiveness of these models and addressing the challenges in cross-lingual hope speech detection, we can pave the way for future advancements in this field [17]. Cross-lingual hope speech detection techniques have been explored in previous research. However, there is still a need for comparative analysis to understand the linguistic features of hope speech in different languages, such as English and Spanish [19]. In the realm of sentiment analysis, "Two-level hope speech detection from tweets" presents an innovative dataset crafted for discerning expressions of hope on social media. It transcends conventional binary sentiment classifications, meticulously parsing between 'Hope' and 'Not Hope', while delving into nuanced categories like 'Generalized Hope', 'Realistic Hope', and 'Unrealistic Hope'. Balouchzahi et al.'s study not only introduces this novel dataset but also dives into the intricate annotation method employed to elevate data quality. Furthermore, the research compares various machine learning models, conclusively highlighting the superiority of contextual models in identifying hope speech. This comprehensive approach, blending machine learning and deep learning techniques [20], establishes a groundbreaking standard for dissecting positive emotional states within digital communications [2]. Burnap et al. [12] examine the impact of online suicide-related communication on vulnerable individuals, highlighting social media platforms as a new interconnected forum for such discussions. Studies show limited evidence of a link between exposure to online suicide content and offline suicidal thoughts. Existing research emphasizes the need for more attention to develop and evaluate online prevention efforts. The Children's Hope Scale measures a child's hope level, aiding in identifying those who may need extra support. This instrument is reliable and valid for use with children, offering a quick way to gauge hope in pediatric psychological research [21]. A study explores automatic detection of positive, hopeful content in social media, focusing on India-Pakistan relations and using YouTube comments as the primary data source. The research aims to create algorithms for language recognition and assess expressions of peace versus war in online conversations to improve online moderation through computational methods [22]. In studying advanced NLP methods for hope speech detection, Ullah et al.'s work is important to note. These techniques may have the potential for adapting to the more subtle task of detecting hope speech, requiring further investigation into their transferability and performance across different linguistic contexts [22, 15].

3. Task Overview

The proposed shared task has two tasks in given work one is binary classification for hs and nhs in Spanish. The second task has data in Spanish and has two sub tasks one is binary classification Hope or Not Hope and multi classification is categorized into Generalized Hope, Realistic Hope, Unrealistic Hope and Not Hope. Similarly the second Task has also tweets in English and has same categories as in Spanish language tweets.

3.1. Task 1

The dataset provided for this collaborative initiative comprises 1400 social media comments in the Spanish language, with an additional 200 comments allocated for validation purposes and 400 for testing purpose are shown in table 1.

Table 1
Task 1 Dataset Split

	Train	Dev	Test
nhs	700	100	199
hs	700	100	198
Total	1400	200	397

3.2. Task 2 Spanish

The task 2 has binary classification and multi class classifications the data in Spanish language described table 2 Binary task Spanish dataset and table 3 multi class.

Table 2
Dataset Distribution

	Train	Dev	Test
Hope	2202	351	378
Not Hope	4701	799	769
Total	6903	1150	1152

Table 3
Task 2 Dataset Split (Multi-class Spanish)

	Train	Dev	Test
Not Hope	4701	799	773
Generalized Hope	1151	186	205
Unrealistic Hope	546	91	96
Realistic Hope	505	74	77
Total	6903	1150	1152

3.3. Task 2 English

The task 2 has also binary classification and multi class classifications the data in English language table 4 for English binary dataset and table 5 for in English.

Table 4
Task 2 Dataset Split (Binary English)

	Train	Dev	Test
Hope	3104	530	527
Not Hope	3088	502	484
Total	6192	1032	1032

Table 5
Task 2 Dataset Split (Multi-class English)

	Train	Dev	Test
Not Hope	3088	502	489
Generalized Hope	1726	300	301
Unrealistic Hope	648	102	106
Realistic Hope	730	128	120
Total	6192	1032	1032

4. Methodology

The proposed methodology comprises two main stages: Preprocessing and Model for classifications, aimed at classifying text into two categories: "hs", "nhs" for task 1 and for task two binary classification and multi classifications "Not Hope", "Generalized Hope", "Unrealistic Hope", "Realistic Hope" for Spanish and English languages.

4.1. Preprocessing

Preprocessing involves cleansing data to eliminate noise, thereby enhancing data quality and improving performance (This entails removing punctuation symbols, numerical data, commonly occurring words, stopwords and uninformative phrases (such as those starting with @), as they don't add value to the classification task. Additionally, uppercase characters in Latin script are converted to lowercase to reduce the number of distinct words. Also we remove urls, emojis and empty rows and we get only that has our desired labels for both tasks.

4.2. Model Description

In this work we used different models but we mention only best model BERT.Our task aimed at identifying hope speech, not hope and multi class classification. We employed the BERT (Bidirectional Encoder Representations from Transformers) multilingual Transformer model, which is well-known for its contextual comprehension of text across different languages. Our approach involved extensive experimentation to fully utilize the capabilities of this advanced technology within our domain. Our primary objective was to develop a robust and accurate system for detecting and categorizing hate speech. By leveraging the BERT multilingual Transformer model, we aimed to create a highly capable system capable of effectively recognizing and classifying hope speech content. We extensively explored and experimented with this model to identify the most optimal architecture and configurations that would result in superior performance in identifying and mitigating hope and not hope and different hopes categorical content in textual data. This process included fine-tuning the model parameters, experimenting with various training methodologies, and optimizing the model's ability to understand and categorize hope expressions. Our ultimate goal was to achieve heightened accuracy and efficiency in the detection and classification process. By utilizing the BERT multilingual Transformer, we endeavored to harness cutting-edge technology and explore its potential to improve the effectiveness of hate speech identification systems through state-of-the-art natural language understanding and classification capabilities.

4.3. Dataset Split

The dataset consists of a training set and a validation set, with a portion of the labeled data allocated for training the BERT multilingual model and the remaining portion reserved for validation. This substantial portion serves as the foundation for the model to learn and extract patterns, linguistic nuances, and indicators of not hope speech from the provided for both tasks. The model undergoes the

training process using this data to adjust its parameters and optimize its understanding of not hopeful expressions.

Simultaneously, a smaller subset, constituting of the labeled dataset, is set aside as the validation set. This portion is crucial for fine-tuning the model's performance and validating its effectiveness. The validation set assists in adjusting hyperparameters, evaluating the model's performance on unseen data, and preventing overfitting, ultimately enhancing the model's generalization. It provides a means to measure how well the model learns from the training data and how effectively it can predict instances of not hope speech in new, unseen instances of test data.

Finally, the unlabeled test data, separate from the training and validation sets, serves as a means to assess the model's real-world performance. This dataset, containing instances of text without labeled categories, enables the evaluation of how well the trained BERT multilingual model can generalize its learning and accurately classify instances of not hope speech.

4.4. Model Parameters

The provided table delineates key parameters essential for configuring Bert model. It encompasses details such as batch sizes for training, evaluation, and prediction, indicating the number of data points processed simultaneously during these phases. Additionally, it specifies parameters like learning rate, the number of training epochs, and the maximum sequence length, crucial for model training and performance. Furthermore, it includes parameters such as dropout rate, training resolver (optimizer), and loss function, which are fundamental components in shaping the model's architecture and training process. This comprehensive set of parameters serves as a foundation for fine-tuning the model's behavior and optimizing its performance across various tasks.

Table 6
Model Configuration Parameters

Parameter	Value
TRAIN_BATCH_SIZE	32
EVAL_BATCH_SIZE	8
PREDICT_BATCH_SIZE	8
LEARNING_RATE	2e-5
NUM_TRAIN_EPOCHS	3.0
MAX_SEQ_LENGTH	128
DROP_OUT	0.1
TRAINING_RESOLVER	Adam
LOSS_FUNCTION	Binary Cross-entropy

5. Results and Discussions

The provided table displays the performance metrics of a model across various classification tasks. Each row corresponds to a specific task, with columns representing different evaluation metrics. The "M_Pr", "M_Re", and "M_F1" columns indicate precision, recall, and F1-score respectively, computed using a macro-average approach. Similarly, the "W_Pr", "W_Re", and "W_F1" columns represent precision, recall, and F1-score, but weighted by class frequency.

For instance, in the "EDI" task, the model achieves a precision of approximately 0.59, recall of 0.58, and an F1-score of 0.57. These metrics suggest a balanced performance across precision and recall, with the F1-score reflecting the harmonic mean of the two. Notably, the binary classification tasks in Spanish and English exhibit higher precision and recall compared to the multiclass tasks. The binary Spanish task particularly stands out with an F1-score of 0.71, indicating robust performance in distinguishing between classes.

Conversely, the multiclass tasks in both Spanish and English demonstrate more varied performance, with precision, recall, and F1-score values differing across the tasks. These discrepancies could stem from various factors such as class imbalances, data quality, or the complexity of distinguishing between multiple classes. Overall, the table provides a comprehensive overview of the model’s performance across different classification tasks, aiding in the assessment and refinement of the model’s capabilities.

Table 7
Task Results

Tasks	M_Pr	M_Re	M_F1	W_Pr	W_Re	W_F1
EDI	0.59	0.58	0.57	0.59	0.58	0.57
Binary Spanish	0.71	0.72	0.71	0.75	0.74	0.74
Multi class Spanish	0.47	0.30	0.30	0.60	0.66	0.59
Binary English	0.74	0.74	0.74	0.74	0.74	0.74
Multi class English	0.54	0.44	0.45	0.58	0.59	0.56

6. Conclusion

In conclusion, our study focused on detecting hopeful communication through the HOPE: Multilingual Hope Speech Detection challenge. While achieving a promising macro F1 Score of 0.74 for English tweets and 0.71 for Spanish. Spanish text detection remains a challenge. Our research highlights the significance of understanding hope speech in the digital realm, aiding in sentiment analysis and fostering positivity online. As social media platforms shape public discourse, identifying hope speech provides insights into societal interactions. Despite challenges posed by negative content, promoting an optimistic online environment is vital for mental well-being. Continued research is crucial for mitigating harmful content and nurturing positive online communities.

7. Acknowledgments

The work was done with partial support from the Mexican Government through the grant A1- S-47854 of CONACYT, Mexico, grants 20241816, 20241819, and 20240951 of the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, Mexico. The authors thank the CONACYT for the computing resources brought to them through the Plataforma de Aprendizaje Profundo para Tecnologías del Lenguaje of the Laboratorio de Supercómputo of the INAOE, Mexico and acknowledge the support of Microsoft through the Microsoft Latin America PhD Award.

References

- [1] 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.
- [2] F. Balouchzahi, G. Sidorov, A. Gelbukh, Polyhope: Two-level hope speech detection from tweets, *Expert Systems with Applications* 225 (2023) 120078.
- [3] F. Balouchzahi, G. Sidorov, A. Gelbukh, Polyhope: Two-level hope speech detection from tweets, *Expert Systems with Applications* 225 (2023) 120078. doi:10.1016/j.eswa.2023.120078.
- [4] 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).
- [5] B. R. Chakravarthi, Hope speech detection in youtube comments, *Social Network Analysis and Mining* 12 (2022) 75.

- [6] S. Dowlagar, R. Mamidi, Edione@ It-edi-eacl2021: Pre-trained transformers with convolutional neural networks for hope speech detection., in: Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion, 2021, pp. 86–91.
- [7] M. Zamir, M. Tash, Z. Ahani, A. Gelbukh, G. Sidorov, Tayyab@ dravidianlangtech 2024: detecting fake news in malayalam lstm approach and challenges, in: Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, 2024, pp. 113–118.
- [8] E. Diener, The science of well-being: The collected works of Ed Diener, volume 37, Springer Science & Business Media, 2009.
- [9] M. Zamir, M. Tash, Z. Ahani, A. Gelbukh, G. Sidorov, Lidoma@ dravidianlangtech 2024: Identifying hate speech in telugu code-mixed: A bert multilingual, in: Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, 2024, pp. 101–106.
- [10] 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.
- [11] 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.
- [12] P. Burnap, G. Colombo, R. Amery, A. Hodorog, J. Scourfield, *Online social networks and media* (2017).
- [13] D. N. Milne, G. Pink, B. Hachey, R. A. Calvo, Clpsych 2016 shared task: Triaging content in online peer-support forums, in: Proceedings of the third workshop on computational linguistics and clinical psychology, 2016, pp. 118–127.
- [14] T. Elmer, K. Mephram, C. Stadtfeld, Students under lockdown: Comparisons of students' social networks and mental health before and during the covid-19 crisis in switzerland, *Plos one* 15 (2020) e0236337.
- [15] K. Puranik, A. Hande, R. Priyadharshini, S. Thavareesan, B. R. Chakravarthi, Iiitt@ It-edi-eacl2021-hope speech detection: there is always hope in transformers, *arXiv preprint arXiv:2104.09066* (2021).
- [16] B. R. Chakravarthi, Multilingual hope speech detection in english and dravidian languages, *International Journal of Data Science and Analytics* 14 (2022) 389–406.
- [17] B. R. Chakravarthi, B. Bharathi, J. P. Mccrae, M. Zarrouk, K. Bali, P. Buitelaar, Proceedings of the second workshop on language technology for equality, diversity and inclusion, in: Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion, 2022.
- [18] 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.
- [19] 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.
- [20] A. A. Eponon, I. Batyrshin, G. Sidorov, Pinealai_stressident_It-edi@ eacl2024: Minimal configurations for stress identification in tamil and telugu, in: Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity, Inclusion, 2024, pp. 152–156.
- [21] C. R. Snyder, B. Hoza, W. E. Pelham, M. Rapoff, L. Ware, M. Danovsky, L. Highberger, H. Ribinstein, K. J. Stahl, The development and validation of the children's hope scale, *Journal of pediatric psychology* 22 (1997) 399–421.
- [22] 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).