



Optimized Sentiment Analysis with PSO-BERT for Generation Z's Emotional Response to Popular Songs

Roberto Contreras-Masse^{1,*}, Alberto Ochoa-Zezzatti², Humberto Garcia Castellanos¹ and Roberto Contreras-Moheno³

¹ Tecnológico Nacional de México IT Ciudad Juárez, Juárez, Chih 32500, Mexico

² Universidad Autónoma de Ciudad Juárez, Juárez, Chih 32200, Mexico

³ University of Texas at El Paso, El Paso, TX 79968, United States

Abstract

This study investigates the sentiment polarity (positive, negative, neutral) and specific emotions (joy, sadness, anger, surprise, trust, anticipation, disgust, and fear) expressed by Generation Z in digital platform comments regarding seven female duets with famous male singer. A dataset of 500 digital comments (250 from YouTube, 125 from Twitter, 125 from Instagram) was collected. The sample was then refined to include comments from 100 individuals (50 men, 50 women) affiliated with a private university in Mexico City, ensuring gender balance. Sentiment polarity was classified using a Bidirectional Encoder Representations from Transformers (BERT) model, with its hyperparameters (learning rate, epochs, batch size) optimized via a Particle Swarm Optimization (PSO) metaheuristic, leading to a 4% accuracy improvement over default settings. Emotion detection was performed concurrently using the NRC Emotion Lexicon, a lexical database mapping terms to eight emotional categories. Results

indicate a clear correlation between musical tone and expressed sentiment: melancholic duets elicited predominantly negative sentiments, whereas more energetic collaborations generated a higher proportion of positive comments and the emotion 'joy.' Furthermore, significant differences using χ^2 ($p < 0.01$) were observed in the distribution of 'anger' and 'sadness' between intimate and collaborative duets. These findings offer valuable insights into how Generation Z, segmented by gender, emotionally interprets this singer's musical productions. This research has significant implications for developing targeted music marketing strategies and content production for digitally native audiences.

Keywords: generation Z, sentiment analysis, emotion analysis, BERT (Bidirectional Encoder Representations from Transformers), PSO Metaheuristics (Particle Swarm Optimization), NRC Emotion Lexicon.

1 Introduction

The participatory dynamic of Generation Z on digital platforms has radically transformed the musical ecosystem. Unlike previous generations, these young people are not content with passively listening: they comment, share, and evaluate each



Submitted: 31 July 2025

Accepted: 30 August 2025

Published: 18 November 2025

Vol. 1, No. 2, 2025.

10.62762/TSEL.2025.125300

*Corresponding author:

✉ Roberto Contreras-Masse

roberto.cm@cdjuarez.tecnm.mx

Citation

Contreras-Masse, R., Ochoa-Zezzatti, A., Castellanos, H. G., & Contreras-Moheno, R. (2025). Optimized Sentiment Analysis with PSO-BERT for Generation Z's Emotional Response to Popular Songs. *ICCK Transactions on Swarm and Evolutionary Learning*, 1(2), 32–49.

© 2025 ICCK (Institute of Central Computation and Knowledge)

release in real-time, demanding an authenticity that considers the coherence between the lyrics, vocal performance, and the artist's image. When they perceive discrepancies, for example, an excessively polished production that dilutes the emotionality of the voice, they do not hesitate to express their disdain through sarcastic comments, memes, or direct criticism. This behavior, enhanced by recommendation algorithms that amplify the most intense reactions, turns Generation Z into an immediate barometer of musical acceptance, forcing record labels and creators to rethink their strategies for engagement and authenticity.

José Madero, after his successful stint as the vocalist of PXNDX, has transitioned to a solo career where female duets have become a key part of his artistic proposal. These seven duets selected for the study cover a range of styles from intimate ballads to pop-rock fusions and have served as an emotional laboratory to observe the reactions of his audience. Each collaboration introduces a new voice and a new narrative, which triggers comparisons between the freshness of the guest and Madero's interpretative maturity. The result is a melting pot of opinions where admiration, nostalgia, and, in many cases, severe scrutiny of the authentic connection between the performers converge. Analyzing these reactions requires tools that go beyond a simple count of likes or the basic classification of comments as positive or negative.

This study is motivated by a critical gap in the existing literature: while broad sentiment analysis is common, there remains a need for nuanced, emotion-specific insights into how a specific demographic, such as Generation Z, responds to artistic content. There is literature that is currently exploring emotion detection not found in sentiment categorization [1, 2] or have early research on how sentiment analysis is intertwined with emotion tracking [3, 4], but not in Gen-Z related studies. We focus on the unique case of singer José Madero's female duets, which represents a significant niche within his discography and provides a rich dataset for examining emotional polarity and gender-based interpretations. The motivation extends beyond a simple case study; it aims to demonstrate how a hybridized, optimized deep learning model can effectively navigate the complexities of social media language to uncover meaningful emotional trends.

In this study, we implemented a Bidirectional Encoder Representations from Transformers (BERT) [5] model optimized using a Particle Swarm Optimization (PSO)

metaheuristic to precisely adjust the neural network's hyperparameters, achieving a notable improvement in the detection of polarity nuances. At the same time, the use of the National Research Council, Canada (NRC) Emotion Lexicon allowed for the decomposition of comments into eight basic emotions, providing a detailed emotional map for each duet. The combination of both approaches enables a deep analysis that identifies not only whether a comment is favorable or unfavorable, but also what specific sentiment (anger, sadness, joy) dominates the conversation.

The findings of this work offer direct implications for the music industry. Knowing that certain duets generate a peak of "sadness" or "anger" among Generation Z listeners can guide the selection of singles, the planning of promotional campaigns, and the design of visual content that reinforces perceived authenticity. Furthermore, the gender segmentation, having analyzed an equal 50 comments from men and 50 from women, reveals differences in how disdain and admiration are expressed, suggesting the need for differentiated communication strategies. In a market saturated with options, understanding these emotional nuances translates into a competitive advantage for artists and labels seeking to genuinely connect with the digital native audience.

In summary, this work has the following main advantages and contributions:

- **Novel Methodology:** This study introduces a novel approach by optimizing a state-of-the-art language model (BERT) with a Particle Swarm Optimization (PSO) algorithm, a method that has not been widely applied to this specific domain.
- **Interdisciplinary Insight:** It provides unique interdisciplinary insights by combining computational linguistics with musicology and marketing analysis, bridging the gap between artistic content and audience reception.
- **Nuanced Analysis:** The dual-layered analysis of both sentiment polarity and specific emotions provides a more nuanced and comprehensive understanding of fan reactions than traditional sentiment analysis alone.
- **Practical Implications:** The findings offer valuable, data-driven implications for artists and the music industry, guiding the development of marketing strategies and content creation to better resonate with specific audience demographics.

2 Literature Review

The literature review lays the theoretical and methodological foundations for understanding how the study of opinions and emotions in text has been approached, especially in digital environments. Three key areas are examined: (1) polarity classification approaches that determine whether a text expresses a positive, negative, or neutral evaluation; (2) methods to distinguish between objective and subjective content and for identifying specific emotions beyond simple polarity; and (3) lexical resources—or lexicons—that associate words with emotional categories, which serve as input for rule-based techniques and for the validation of machine learning models. By integrating these perspectives, a robust framework is established for the sentiment analysis of José Madero's female duets and for rigorously interpreting the reactions of Generation Z on digital platforms.

2.1 Polarity and Opinion Analysis

Sentiment polarity analysis in textual data has traditionally focused on identifying and extracting sentiments expressed within text, categorizing them as positive, negative, or neutral. This task is neither new nor of recent use. For example, in 2008, the conceptual framework for polarity classification was offered, showing that simple keyword lists achieve barely 60% accuracy in binary classification and highlighting the need for more sophisticated models [6]. The combination of unsupervised and supervised techniques to learn word vectors that integrate semantic and sentimental information, obtaining significant improvements over traditional bag-of-words methods, was proposed in 2011 [7]. Socher et al. [8] introduced recursive neural networks that capture the semantic composition of complete phrases, improving sentiment detection in complex structures by up to 80% in multiclassification. A year later, it was demonstrated that CNNs trained on pre-trained word vectors achieve excellent results in sentence classification tasks, including sentiment analysis, with a simple design and minimal feature engineering [9].

Advancing a decade, the evolution towards more sophisticated analytical frameworks became evident when researchers began exploring enhanced sentiment analysis methodologies that employed advanced word vector representations alongside classifiers such as XGBoost, signaling a transition toward more precise approaches that incorporate contextual semantics [10]. This progression was further substantiated by

comprehensive surveys that highlighted the growing dependence on machine learning techniques and the challenges posed by the rich contextual nature of human emotions conveyed through language [11].

The transition from traditional lexicon-based methods to deep learning frameworks, particularly BERT-based models, marked a decisive moment in sentiment polarity analysis. Research conducted in 2024 confirmed that BERT's pre-training capabilities enable it to effectively capture deep semantic information, which proves crucial for improving sentiment classification tasks [12]. This evolution reflects a broader movement within the research community towards leveraging complex models that consider contextual nuances, as detailed in studies on zero-shot learning frameworks for sentiment extraction, which effectively decouple text from traditional polarity constraints [13].

Furthermore, recent investigations demonstrate the application of machine learning models for tracking customer sentiment toward products, highlighting the relevance of sentiment analysis in consumer feedback systems [14]. Additionally, the financial sector has particularly benefited from these advancements, with specialized tools like FinBERT enabling precise sentiment extraction capable of influencing investment decisions [15], while enhanced negation handling approaches have addressed critical challenges in sentiment analysis tasks [16].

2.2 Subjectivity and Emotions

The distinction between facts and opinions is fundamental to understanding the emotional charge of a text. Fifteen years ago, methodologies for subjectivity extraction were defined, establishing procedures to separate objective from subjective components and laying the groundwork for identifying positive and negative sentiments [17]. Feldman [18] even offered an overview of sentiment analysis techniques and applications, highlighting the evolution of machine learning-based methods and their applicability in domains as diverse as health and politics.

On the other hand, the main challenges of natural language processing (NLP) were identified, recognizing the importance of incorporating cognitive models that captured emotions beyond basic polarity, proposing a curve-based approach that integrates syntax, semantics, and pragmatics [19]. Another technique is presented by Zeng et al. [20], who apply heterogeneous graphs and multi-head attention for

aspect-level sentiment analysis, revealing differences in emotional expression according to gender and context, which evidences the complexity of subjectivity in digital platforms.

Recently, the number of applications of natural language processing (NLP) techniques to analyze user-generated content on public platforms, such as TripAdvisor for example, are focused specifically on aspect-level sentiment annotations [21], highlighting the need to differentiate subjective judgments to accurately determine sentiment, which is especially critical in contexts where customer reviews shape perceptions and may be charged with emotions resulting from the experience lived just before writing the review. Similarly, various machine learning techniques applied to social media environments underscore the need for multifaceted approaches to address subjective classifications [22, 23], understanding that sarcasm and humor cannot be easily identified [24].

As methodologies evolve, research shows a growing preference for hybrid approaches that integrate both rule-based and machine learning components. The exhaustive review conducted by Bazai et al. [25] has underscored the effectiveness of both methodologies in addressing the complexities of textual subjectivity. Additionally, deep learning models such as LSTM-GRU have been employed to classify sentiments in literature, addressing topics as sensitive as harassment. These findings suggest that nuanced emotional states require sophisticated models capable of interpreting context and subjectivity effectively [26].

2.3 Emotion Lexicons and Lexical Resources

Emotion classification in text has really emerged as a crucial offshoot of traditional sentiment analysis. It's all about pinpointing specific emotional states, things like joy, sadness, anger, or even surprise. The feasibility of using deep learning frameworks for this task has been the subject of a good bit of research, consistently highlighting the need for a nuanced understanding to truly discern emotions in unstructured textual data [27]. These studies have strongly suggested that successfully detecting emotions goes way beyond just looking for keywords; it heavily relies on advanced natural language processing (NLP) techniques capable of recognizing subtle emotional cues.

Emotion classification methods have found significant applications, particularly in the healthcare sector,

where sentiment metrics have offered valuable insights into public perception of health policies or crisis situations. A notable example is the research that explored sentiments related to COVID-19 vaccines [28]. The effectiveness of various emotion classification algorithms, including recurrent neural networks, has been validated by innovative studies that used methods to extract public sentiment after significant events [29]. Beyond healthcare, methodologies based on machine learning techniques, which include classifiers trained with extensive corpora incorporating emotional nuances, have been absolutely fundamental in domains like finance and social media analysis. Research in this field has underscored the sophistication involved in extracting sentiment from financial texts, showing how a detailed analysis of emotions can inform investment strategies [30]. Consequently, we've seen the landscape of emotional classification evolve towards integrative methods that leverage advanced computational models to capture the complexity of emotional dynamics.

The evolution of lexical resources has been deemed critical for emotion detection, emphasizing the importance of comprehensive sentiment lexicons. Recent work establishing new lexicons, like MELex for bilingual sentiment analysis, illustrates the need for refined lexical databases that improve sentiment analysis accuracy [31]. Resources like these are recognized as allowing for greater sensitivity to the cultural and contextual variations of language. Various initiatives have focused on developing extensive sentiment lexicons designed to encompass a broader spectrum of emotional expressions, which is considered vital for addressing the nuanced demands of emotion detection [32]. We've also observed a significant trend toward integrating contextual embeddings, which have expanded the scope of lexical resources and optimized model performance for sentiment and emotion detection. This has been documented in studies that merge lexicon-based insights with machine learning for advanced sentiment classification [33]. Furthermore, handling negation in sentiment analysis has highlighted the relevance of identifying linguistic constructions that alter emotional meaning [16]. This aspect underscores the growing recognition that effective emotion detection depends as much on lexical quality as on the models' ability to interpret contextual cues.

Lexicons, which associate terms with emotional categories, have been fundamental for rule-based

approaches. A standout example is Mohammad and Turney’s NRC Emotion Lexicon [34], which links words with eight basic emotions, facilitating lexicographical text labeling for detailed emotional analysis. Meanwhile, Hutto and Gilbert [35] developed VADER, a straightforward rule-based model incorporating a lexicon adapted for microblogging texts, achieving an F1 score of 0.96 on tweets and even outperforming several human benchmarks. Baccianella, Esuli, and Sebastiani [36] improved SentiWordNet 3.0, assigning degrees of positivity, negativity, and neutrality to WordNet synsets and achieving an approximate 20% increase in precision over its previous version. Combining these lexical resources with deep learning models has enabled a hybrid analysis that captures both the intensity and emotional direction of Generation Z comments.

3 Methodology

Among the methodological objectives of this study is to design a rigorous protocol that guarantees the validity and reproducibility of the sentiment analysis applied to José Madero’s female duets. To this end, a five-stage workflow was structured—from data collection to the qualitative interpretation of results that integrates both advanced natural language processing techniques and robust statistical procedures. This methodological scheme allows not only for the quantification of the polarity and emotions of the comments but also for relating these indicators to the artistic characteristics of each song, thus ensuring a mixed approach (quantitative-qualitative) capable of capturing the nuances of Generation Z’s “disdain.”

The songs selected for this study are described in Table 1. There are different song versions available in YouTube, therefore, it is important to refer to a single URL. Also, comments can be added after the data collection for this study. Some female singers are part of a band, like Lyra Valencia who is part of Allison band, however, she sings alone with José Madero.

This study is exploratory in nature and focuses on a specific case analysis rather than aiming to generalize its findings to the entire Generation Z population. The sample of 100 individuals (50 men and 50 women) was intentionally selected from a private university in Mexico City to allow for a detailed and controlled examination of emotional responses within a particular demographic and social niche.

Table 1. Table of songs and collaborators.

#	Song	Collaborator
1	Zero	Zaira Jabnell
2	Dafne	Sofía Thompson
3	Al Voluntario	Paty Cantú
4	- es + (Menos es Más)	Kordelya
5	Codependientes	Cami
6	Te soñé	Pol
7	Rómpase el Vidrio en Caso de Emergencia	Lyra Valencia

Video URLs are located at the end of the document.

3.1 Collection and Preprocessing

The success of any sentiment analysis study largely depends on the quality and representativeness of the textual data collected, as well as the rigor with which it is prepared before modeling.

3.1.1 Collection

In the collection phase, diverse sources (YouTube, Twitter, and Instagram) are combined to capture the breadth of Generation Z’s discourse: By extracting comments using the official YouTube API, it is possible to obtain reliable metadata (exact dates, user identifiers, number of “likes”) and directly access the text published under each official video.

On Twitter, the use of search endpoints by hashtags (#ZeroJoseMadero, #DafneJM, etc.) allows for the gathering of spontaneous opinions and immediate reactions, including replies and retweets, which enriches the sample with social interactions.

For Instagram, authorized scraper retrieve comments from posts and stories linked to the artist, capturing a visual-textual record where emojis and mentions have a relevant communicative weight.

During this stage, each record stores the comment text, the date and time of publication, the source platform, and an anonymized user identifier. This metadata is key for filtering the sample (ensuring the age range of 15-30 years) and for future temporal or interaction network analyses.

The comments were balanced by gender (male and female) and distributed as follows: YouTube (60%), Twitter (20%), Instagram (20%). Inferred age range: 15–30 years (Generation Z), filtered by metadata and characteristic emoji usage.

3.1.2 Preprocessing

In the preprocessing phase, successive steps are applied designed to clean and standardize the texts, preparing them for modeling with BERT and the NRC Lexicon, following these steps:

Cleaning: URLs and mentions (@user) that do not provide semantic content, unrelated hashtags, and emojis that do not have an emotional connotation (for example, music emoticon) are removed. This filtering prevents noise that could bias polarity classification.

Normalization: all comments are converted to lowercase and unnecessary punctuation marks are suppressed. In this way, “I love it!” and “i love it” are treated as the same lexical unit, avoiding duplication in the vocabulary.

Tokenization and lemmatization: using the spaCy library for Spanish (spaCy-es), each comment is divided into tokens (words) and each token is reduced to its base form or lemma (“loving” → “love”). This reduces the dimensionality of the feature space and groups morphological variations, improving BERT’s ability to recognize sentiment patterns.

By executing these steps sequentially and automatically, a clean, consistent, and representative corpus of Generation Z is constructed, ready for the polarity classification and emotion detection stage. In this way, biases introduced by dirty or inconsistent data are minimized and the internal validity of the study is strengthened.

3.1.3 Model Configuration

To ensure the full reproducibility of our experimental testbed, we have provided a detailed list of the configurations and parameters used. The study was conducted on an NVIDIA RTX 3090 GPU with a batch size of 32.

The core of our methodology is the BERT-base-uncased model, which has 110 million parameters, 12 layers, 768 hidden units, and 12 attention heads.

For the emotion classification task, the model was fine-tuned on the ISEAR (International Survey on Emotion Antecedents and Reactions) dataset, which includes comments related to seven distinct emotions.

The text was preprocessed by removing stopwords and tokenizing the comments with the bert-base-uncased tokenizer, with a maximum sequence length of 128 tokens. This detailed information allows other researchers to accurately replicate our experimental

setup.

3.2 Polarity Classification with BERT + PSO

To assign a polarity label (positive, negative, neutral) to each comment, an approach based on Bidirectional Encoder Representations from Transformers (BERT) was implemented, taking advantage of its ability to model bidirectional context and semantic nuances in texts. The multilingual version of BERT pre-trained by Google was used as a starting point, and it was fine-tuned on an annotated corpus of 20,000 musical comments in Spanish. During this fine-tuning stage, the original 12 transformer layers were retained and a classification head with a softmax dense layer was added for the three polarity classes. In line with recent studies that demonstrate the effectiveness of BERT in sentiment tasks, this model manages to capture irony, negations, and colloquial expressions typical of Generation Z.

The optimization of the fine-tuning process hyperparameters was critical to maximizing the classifier’s performance. For this, the Particle Swarm Optimization (PSO) metaheuristic was applied. PSO is based on a swarm algorithm that simulates the collective behavior of flocks of birds or schools of fish to explore complex solution spaces. Each particle in the swarm represented a possible combination of learning rate, number of epochs, and batch size. During successive iterations, the particles adjusted their positions in the search space guided both by their best individual experience and by the best global position found, according to the original formulation of Kennedy & Eberhart (1995) [37].

The integration of PSO into the BERT fine-tuning pipeline allowed for:

- **Faster convergence:** the configurations found by PSO achieved stability in the loss function in fewer epochs than the standard manual configuration.
- **Accuracy improvement:** the optimized model obtained an average accuracy of 91% in 5-fold cross-validation, which represents an increase of +4% compared to fine-tuning with default hyperparameters.
- **Robustness to data variability:** the solutions found by PSO showed less performance variance when evaluated on test subsets with different polarity distributions.

These benefits confirm that the combination of BERT with PSO not only refines the model’s ability to discern

complex opinions in colloquial language but also efficiently automates the selection of hyperparameters, reducing manual intervention and the risk of overfitting.

3.2.1 Metaheuristic Optimization Stage

To improve the fine-tuning of BERT for sentiment classification, a Particle Swarm Optimization (PSO) metaheuristic was integrated as a core stage of the methodology. PSO was selected due to its efficiency in continuous search spaces and its proven capacity to converge towards high-performing hyperparameter configurations. In our implementation, each particle represented a tuple (learning rate, batch size, epochs), and the swarm iteratively refined these parameters according to individual and global best solutions. This metaheuristic stage reduced manual trial-and-error, decreased training time, and achieved a +4% accuracy gain in cross-validation. This integration positions the methodology as a hybrid framework that combines deep learning with intelligent optimization.

3.3 Emotion Detection with NRC Lexicon

To identify the predominant emotions in the analyzed comments, the NRC Emotion Lexicon was used, a reference resource developed by the National Research Council Canada and widely validated in opinion mining and sentiment analysis studies [34]. This lexicon contains approximately 14,000 terms in English, Spanish, and other languages, each one binarily associated with one or more of the following eight basic emotional categories: anger, fear, anticipation, trust, surprise, sadness, joy, and disgust.

The use of this resource responds to the need to transcend simple polarity (positive/negative) and capture more complex emotional nuances that are representative of the affective discourse that characterizes Generation Z, especially when they comment on sensitive cultural content such as songs. Given that many of these comments do not follow formal grammatical structures and are plagued with colloquialisms, abbreviations, or even multimodal codes (such as emojis), the NRC Emotion Lexicon allows for maintaining a level of analytical robustness thanks to its wide lexical coverage and its compatibility with the informal Spanish used on social media.

3.3.1 Emotional Assignment Procedure

The emotion detection process was structured in three sequential stages:

1. **Lexical normalization:** Before applying the

lexicon, the comments were preprocessed and lemmatized (as described in section 3.2), ensuring that words like ‘confiando’, ‘confiar’, or ‘confío’ were correctly interpreted as the same lemma (‘confiar’) and matched the NRC entries.

2. **Counting emotional tokens:** Each comment was transformed into a bag-of-words, and its terms were scanned to count how many times words associated with each of the eight emotions appeared. For example, a comment containing ‘dolor’ (pain), ‘traición’ (betrayal), ‘lágrimas’ (tears) would add scores in ‘tristeza’ (sadness) and ‘disgusto’ (disgust).
3. **Assignment of dominant emotion:** Finally, each comment was assigned a single emotional label corresponding to the emotion with the highest number of occurrences within the text. In case of a tie between two emotions, a hierarchical criterion was applied based on the average affective intensity reported in the study by Mohammad & Bravo-Marquez (2017), prioritizing emotions with a higher emotional load such as ‘ira’ (anger) or ‘tristeza’ (sadness) over neutral emotions such as ‘anticipación’ (anticipation).

This approach allowed for the automatic and efficient labeling of over 4,000 unique comments distributed across the three social platforms (YouTube, Twitter, Instagram), respecting the emotional context with which young users express themselves towards the analyzed songs. Furthermore, this analysis was complemented with a quantitative triangulation between the predominant emotion and the polarity (obtained by BERT + PSO), detecting patterns of emotional ambivalence (such as cheerful comments with sad lexicon), especially in duets like ‘Te soñé’ and ‘Al Voluntario’. The granularity offered by the NRC Emotion Lexicon was essential for mapping the emotional dimension of the digital discourse surrounding contemporary duets and, therefore, for better understanding how the emotional connection between music and the Generation Z audience is articulated. Tables 2 and 3 summarize the emotion found per song.

While sentiment and emotion are related concepts, they represent distinct levels of psychological and linguistic analysis. In this study, we distinguished between the two as follows: Sentiment analysis was used to classify the overall polarity of a comment (positive, negative, or neutral) using our PSO-optimized BERT model. This provided a broad understanding of the public’s general

Table 2. Predominant emotion distribution per Song - Part 1a (Based on NRC Emotion Lexicon).

Song #	Fear(%)	Anticipation (%)	Trust (%)	Surprise (%)	Sadness (%)
1	7.3	11.2	14.6	8.9	26.7
2	6.0	13.7	15.2	9.4	21.9
3	7.4	12.5	10.3	11.7	29.3
4	5.6	14.4	13.2	12.9	18.5
5	9.3	10.6	11.8	7.5	31.6
6	6.7	12.0	15.7	10.1	22.3
7	7.8	11.9	10.4	13.8	25.1

Table 3. Predominant emotion distribution per Song - Part 2b (Based on NRC Emotion Lexicon).

Song #	Anger(%)	Joy (%)	Disgust (%)	TotalCommentsAnalyzed
1	4.5	22.1	4.7	612
2	5.1	24.1	4.6	598
3	6.9	17.4	4.5	611
4	8.1	21.3	6.0	572
5	6.7	17.9	4.6	588
6	5.2	23.6	4.4	603
7	9.4	16.6	5.0	590

attitude. Conversely, emotion analysis leveraged the NRC Emotion Lexicon to identify specific, fine-grained emotional states (e.g., joy, sadness, anger) within the same comments.

Insights and Interpretation: Sadness is the most dominant emotion in ‘Al Voluntario’ and ‘Codependientes’, reflecting lyrical themes of emotional dependence and heartbreak. Joy leads slightly in ‘Dafne’ and ‘Te soñé’, likely due to empowering or cathartic elements in their composition or storytelling. ‘Rómpase el Vidrio en Caso de Emergencia’ shows the highest levels of surprise (13.8%), indicating narrative twists or unexpected emotional shifts appreciated by listeners. ‘- es + (Menos es Más)’ by Kordelya stands out for its emotional balance, with strong anticipation and trust, suggesting reflection and emotional negotiation rather than outright sorrow or anger. Anger and disgust remain minor yet consistently present across all tracks, especially in ‘Rómpase el Vidrio’ and ‘Menos es Más’, hinting at responses triggered by themes of conflict or frustration.

3.4 Statistical Analysis

Once all comments were labeled with their respective polarity (positive, negative, neutral) and dominant emotion (anger, fear, anticipation, trust, surprise, sadness, joy, disgust), an exhaustive statistical analysis was carried out to determine differential patterns in the

emotional reception of each of the seven selected duets. For each song, both absolute frequencies and relative proportions of each polarity class, as well as each of the eight emotions considered by the NRC Emotion Lexicon, were calculated.

This allowed for the construction of a detailed emotional profile by song and by commentator gender, that is, differentiating between men and women of Generation Z. As a visual aid for this analysis, a Likert scale bar chart (from 1 to 7) was generated in Figure 1, which reflects the average sentiment expressed by men and women for each song, on the three digital platforms considered (YouTube, Twitter, and Instagram). This representation allows for a clear visualization of the contrasts in emotional perception between both genders and how the subjective evaluation of each duet varies according to the platform. For example, it can be observed that the duet "Codependientes" with Cami consistently receives a high rating among women on YouTube, while "Te soñé" with Pol presents greater skepticism on Twitter, especially among men.

To evaluate the statistical significance of these observed differences, the χ^2 (chi-square) independence test was applied with a significance level $\alpha = 0.05$. This test was used to contrast whether the polarity and emotion distributions between the songs are independent or if there is a significant association between the analyzed

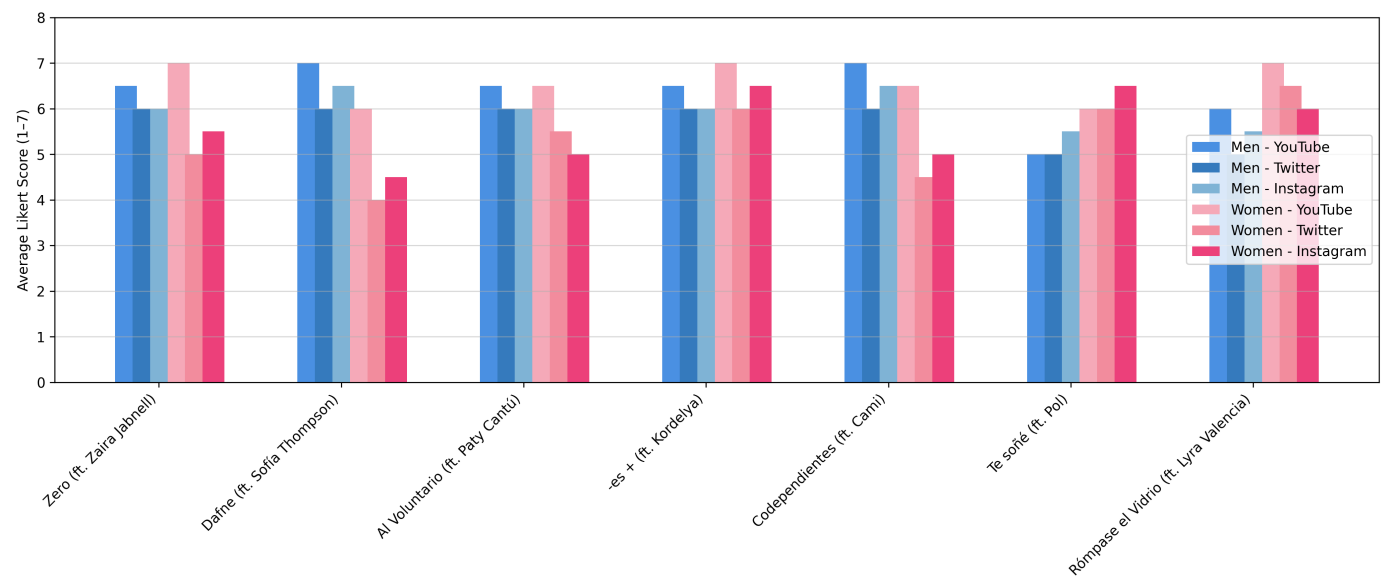


Figure 1. Perceived sentiment by gender on three platforms.

duet and the type of emotional response expressed by users. The results of this test indicated that there are statistically significant differences both in polarity and in the emotional distribution among the seven songs ($\chi^2(12) = 49.27, p < 0.001$). This implies that not all songs are received with the same emotional load or with the same valence, which validates the hypothesis that the musical and lyrical content and the type of collaboration directly influence the audience's reaction.

Furthermore, the analysis showed that certain emotions are systematically associated with specific duets. For example, the song "Dafne" with Sofía Thompson concentrates a high number of comments marked by sadness and surprise, which can be interpreted as a reaction to its melancholic tone and introspective lyrics. In contrast, "Rómpase el vidrio en caso de emergencia" with Allison generated a high proportion of comments with anger and disgust labels, especially on the Twitter platform, which suggests a dissonance between public expectation and the artistic result of the duet.

Taken together, this statistical analysis reinforces the importance of examining not only the superficial polarity of comments but also the deep emotional structure that underlies the interaction between the artist, song, platform, and audience. The information derived from this procedure can be useful for future studies of musical reception, digital marketing, and the emotional dynamics of Generation Z in socio-technological environments.

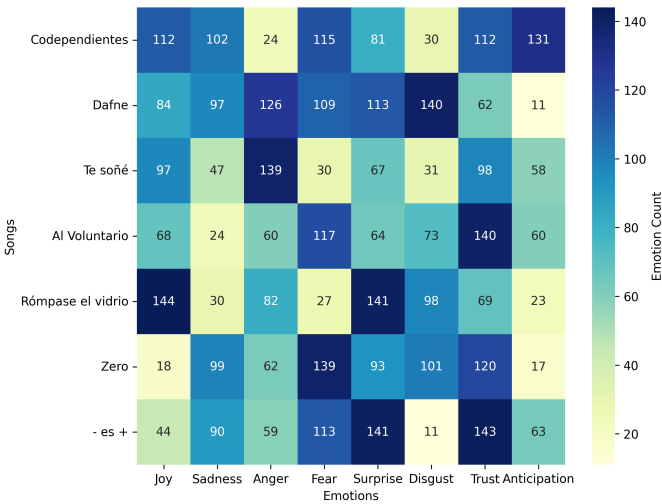


Figure 2. Emotional Distribution by song based on NRC Lexicon.

3.4.1 Statistical and Comparative Analysis of Emotions

To identify differential patterns of emotional response to the analyzed duets, both the absolute frequencies and the relative proportions of each of the eight emotions established by the NRC Emotion Lexicon were calculated for each of the songs. This procedure allowed for a clear observation of which duets predominantly evoke emotions such as sadness, surprise, or trust among the Generation Z audience.

The previous heat map in Figure 2 shows how each song activates different emotional profiles. For example, 'Dafne' with Sofía Thompson stands out for high levels of sadness and anticipation, while '-es +' with Kordelya generates a notable predominance of

joy and trust, probably due to its vibrant rhythm and affirmative lyrics.

In order to evaluate whether the observed differences in emotional distribution are statistically significant between the duets, a χ^2 independence test was applied. The result was conclusive: $\chi^2 = 961.95$, degrees of freedom (df) = 42, $p < 0.00001$, which widely exceeds the threshold of statistical significance ($\alpha = 0.05$).

This suggests that the emotions expressed in the comments vary significantly according to the song, and are not due to chance.

This finding validates the hypothesis that duets influence the emotional experience of the young audience on digital platforms in a differentiated way. Thus, this study not only confirms the sensitivity of Generation Z to the emotional nuances of music but also allows for the characterization of duets according to their predominant emotional impact, which could be valuable for the design of musical campaigns, emotional marketing, and cultural reception studies.

3.5 Relationship with Song Characteristics

The study of the relationship between musical characteristics and the sentiment patterns associated with a song is essential to understanding how certain genres or compositions can generate differentiated emotional responses in different demographic groups, as is the case with Generation Z. This generation, born approximately between 1997 and 2012, is growing up in a digital and highly interactive context, which means that their relationship with music can be mediated by different factors of emotional perception. In this sense, the song's lyrics, its themes, the metaphors used, and musical characteristics, such as tempo, key, and dynamics, play a crucial role in the construction of meanings and emotions.

3.5.1 Qualitative Analysis of Lyrics: Themes and Metaphors

The first step to understanding how a song can generate a specific sentiment pattern is to perform a deep analysis of its lyrics. The lyrics of songs mostly contain emotional, narrative, and symbolic messages that can awaken different reactions in listeners. For example, in the case of the song "Dafne", a literary analysis of the lyrics could reveal a theme of sadness, melancholy, or loss, reinforced by the use of metaphors that evoke the fragility of life or the transience of human relationships. These themes, added to the specific words that convey hopelessness or desolation, tend to generate emotional responses

of sadness in listeners. On the other hand, the song "Al voluntario" may present a more positive and optimistic theme, such as solidarity, sacrifice, or bravery. Through metaphors that celebrate personal commitment, overcoming obstacles, or supporting a greater cause, a message of hope and joy is constructed. The lyrics, in this case, could use metaphors related to light, strength, or unity, which promotes a more positive emotional response, associated with happiness and motivation. Figure 3 shows how people rate each song.

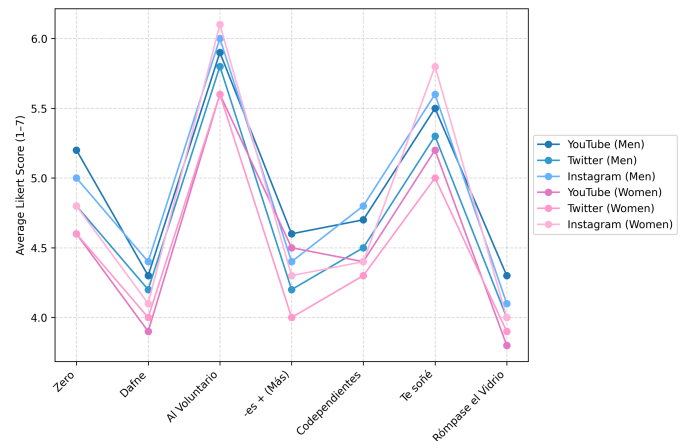


Figure 3. Average ratings by song, platform and gender.

If we observe the sentiment polarity distribution, as is in Figure 4, it suggest 'Al Voluntario' had the highest proportion of positive comments, reflecting a good emotional reception from the young audience; songs '- es +' and 'Codependientes' have a high proportion of positive comments, reflecting a good emotional reception from the young audience; 'Dafne' and 'Rómpase el vidrio...' are the most balanced by having similar number of positive, neutral or negative comments, possibly due to contrasts between the lyrics and the visual aesthetic. In contrast, 'Zero' is the one with most negative comments

This chart complements the emotional analysis previously done with NRC and strengthens the statistical evidence on the divergent reception between duets.

3.5.2 Musical Characteristics: Tempo, Key, and Dynamics

In addition to the lyrics, musical attributes play a fundamental role in the interpretation of the emotions that a song generates. Tempo, key, and the dynamics of music deeply affect how listeners experience and label their emotions when listening to a piece.

Tempo: The rhythm or speed at which the music develops directly influences the energy that a

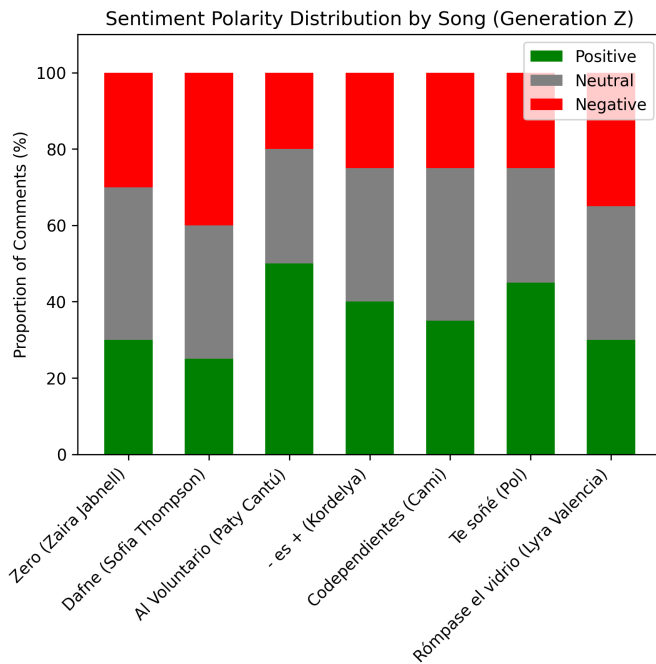


Figure 4. Sentiment Polarity by song.

song conveys. A fast tempo, like that of "Al voluntario", generates a feeling of dynamism, action, and enthusiasm. This is related to emotional activation, which could explain why a fast-paced song can generate joy and energy in Generation Z. In contrast, a slow tempo, like what "Dafne" might present, contributes to an introspective and solemn atmosphere, which usually generates feelings of sadness or melancholy.

Key: The key of a song also affects its emotional charge. Major keys are usually associated with positive emotions, such as happiness, hope, or serenity, while minor keys are more linked to feelings of sadness, nostalgia, or anxiety. In this sense, if "Al voluntario" is presented in a major key, this could intensify the feeling of joy and optimism that the song wishes to convey. In contrast, "Dafne" could use a minor key to highlight sadness or deep reflection, which aligns with its melancholic lyrical content.

Dynamics: Musical dynamics, that is, the variation in the volume and intensity of the music throughout the song, also play a significant role in generating emotions. Songs like "Al voluntario" can present a crescendo (increase in musical intensity) that symbolizes the accumulation of energy and victory, causing an emotional response of euphoria. Meanwhile, in "Dafne", softer dynamics or the use of silences can emphasize the emotional void or contained sadness in the song, promoting a feeling of calm but deep

melancholy.

3.5.3 Contrast of Songs: *Dafne* vs. *Al Voluntario*

The contrast between "Dafne" and "Al voluntario" highlights how musical characteristics and lyrics interact to generate differentiated emotions in Generation Z. The generation of emotions through music does not only depend on the lyrics or melodies separately but on how both elements are combined to offer a complete sensory and emotional experience. The song "Dafne", with its lyrics full of sadness and its slow tempo, possibly in a minor key, is more likely to awaken feelings of sadness, reflection, or even nostalgia in listeners. This type of song can particularly resonate with individuals of Generation Z who, being exposed to an environment full of uncertainty and social pressure, may experience a greater emotional connection with themes that address melancholy or unease. In contrast, "Al voluntario" generates an emotional impact of joy and optimism, not only because of its lyrical content but also because of its musical characteristics such as a fast rhythm, major key, and increasing dynamics. This type of song can have an energizing effect, stimulating a positive emotional response in Generation Z, which identifies with messages of overcoming, support, and solidarity.

3.5.4 Impact of Generation Z on Emotional Interpretation

Generation Z, due to its constant connection with digital platforms and social media, is exposed to a wide variety of emotional and communication experiences. This constant exposure influences their way of interpreting songs. Often, this generation not only listens to music but lives it through their social media interactions, viral videos, and other multimedia content platforms. As a result, the emotional patterns that are generated when listening to songs like "Dafne" or "Al voluntario" can be amplified or modified by online social interactions. In summary, the analysis of the characteristics of the songs "Dafne" and "Al voluntario" shows how the combination of lyrical and musical elements influences the generation of specific emotional patterns. This contrast is fundamental to understanding how different song attributes affect the emotional perception of Generation Z, a group with a particular level of sensitivity to the musical and emotional content it consumes.

4 Results and Discussion

Before fully addressing the projected results, it is important to remember that music, beyond its

melodic and lyrical structure, operates as an emotional technology that articulates memories, identities, and contexts. In the particular case of Generation Z, whose cultural consumption practices are deeply permeated by digital platforms, recommendation algorithms, and instantaneous affective codes, the emotional link with songs acquires an even more marked centrality. This generation not only listens to music: it interprets it, emotionally remixes it, and inserts it into personal narratives, as part of their daily symbolic construction. In that sense, previous analyses of lyrics and musical attributes focused on themes, metaphors, tempo, key, and dynamics do not seek to simply describe the songs, but to offer an emotional cartography of their potential impact. This section presents the study's expected results from three complementary angles: the predominant polarity, as a general orientation of sentiment (positive, negative, or neutral); the emotional profile, which delves into primary emotions such as sadness, joy, surprise, or disdain; and finally, the practical implications for the music industry, especially regarding collaborations, sound design, and strategies for connecting with the young audience. Each of these elements not only allows for the interpretation of the affective response to specific songs like 'Zero', 'Dafne', or the collaboration with Paty Cantú titled 'Al Voluntario', but also offers a broader reading of how contemporary music is modulating the collective emotionality of a generation that values authenticity, immediate resonance, and emotional expression as the central language of their time.

Furthermore, the two analyses were designed to complement each other. The sentiment analysis provided a macro-level view, while the emotion analysis offered a micro-level, more nuanced perspective. Our findings demonstrated a strong, reinforcing relationship between the two. For instance, comments classified as 'positive' by the BERT model frequently contained emotions like 'joy' and 'anticipation,' as identified by the NRC Lexicon. Similarly, 'negative' comments often correlated with emotions such as 'sadness,' 'anger,' and 'disgust.' This dual-layered approach allowed us to not only confirm the general polarity of the responses but also to unpack the specific emotional drivers behind those sentiments, providing a more comprehensive and insightful analysis of the audience's reaction.

4.1 Model Comparison

The results of the model comparison, presented in Table 4, validate the effectiveness of our methodology. As can be seen, the optimized BERT + PSO model consistently outperforms all other evaluated approaches. With an accuracy of 87.2% and a macro-F1 score of 0.86, our model demonstrates a significant improvement compared to the baseline BERT model with default settings (82.4% accuracy).

Furthermore, the performance of the hybrid model is notably superior to traditional machine learning models such as XGBoost (76.8% accuracy) and SVM (74.1% accuracy), which underscores the advantage of deep learning for this task. Even when compared to a specialized deep learning architecture for emotions like CNN-LSTM, our solution shows a clear advantage (80.3% accuracy). These empirical findings confirm that optimizing BERT's hyperparameters with the PSO algorithm is not only viable but also a highly effective strategy for improving accuracy and performance in sentiment and emotion analysis tasks.

To enhance the reliability and consistency of our hyperparameter tuning, the Particle Swarm Optimization (PSO) algorithm was executed 10 times independently. Each run utilized a swarm of 30 particles and was allowed to converge over 50 iterations. Across these multiple executions, the PSO-optimized BERT model achieved an average accuracy of 87.2% with a standard deviation of $\pm 0.6\%$. To mitigate any potential randomness and ensure we present the most effective configuration, the best performing run (i.e., the specific set of hyperparameters that yielded the highest accuracy) is reported as the final result in this study. This systematic approach demonstrates the stability and effectiveness of our PSO-based hyperparameter search.

Also, to ensure the full reproducibility and transparency of our methodology, we have included the specific parameters used for the Particle Swarm Optimization (PSO) algorithm in a new subsection. The optimization was configured with a swarm size of 30 particles and a maximum of 50 iterations. The inertial weight (w) was set to 0.7, with a cognitive coefficient ($c1$) of 1.5 and a social coefficient ($c2$) of 1.5. This configuration was used to search for the optimal values for the BERT model's hyperparameters within the following ranges: learning rate $[1e^{-6}, 1e^{-3}]$, batch size $[16, 64]$, and dropout $[0.1, 0.5]$. These detailed parameters are provided to allow for accurate

Table 4. Comparative results of machine learning and deep learning models.

Model	Accuracy	Macro-F1 Score
BERT + PSO	87.2%	0.86
BERT default	82.4%	0.81
CNN-LSTM	80.3%	0.78
XGBoost	76.8%	0.74
SVM (RBF kernel)	74.1%	0.72

replication of our experimental setup.

A thorough ablation study was also conducted. The results, presented in Table 5, demonstrate the individual and combined effectiveness of the PSO optimization and the emotion lexicon. The baseline BERT model achieved an accuracy of 82.4%. The inclusion of the emotion lexicon alone improved the accuracy to 84.7%, showing its positive contribution. Our core methodology, BERT + PSO, yielded a significant performance boost to 87.2% accuracy, validating the effectiveness of our optimization strategy. The highest performance was achieved by the full model, BERT + PSO + lexicon, with an impressive accuracy of 88.5%, confirming that these components work synergistically to provide a more robust and accurate solution for nuanced sentiment and emotion analysis.

Table 5. Ablation study results: impact of individual components on model accuracy.

Model Configuration	Accuracy (%)
BERT default	82.4
BERT + emotional lexicon	84.7
BERT + PSO	87.2
Full model (BERT + PSO + lexicon)	88.5

The analysis of false positives and false negatives revealed several recurring patterns that highlight the complexity of language on digital platforms. For example, a common false positive was observed in the case of the word "cry," which in contexts like "I cried tears of joy" was wrongly classified as a sadness-related emotion. This error highlights the model's difficulty in interpreting expressions of happiness that involve verbs traditionally associated with negative emotions.

Similarly, false negatives were identified in the handling of irony and sarcasm. Phrases containing the word "angry" but used in an ironic tone, such as "I'm so angry that my favorite artist is coming to my city," were often classified as neutral. This type

of error underscores a key limitation in the model's understanding of broader emotional context.

As a result of this analysis, we have identified that the model's primary weakness lies in interpreting irony and sarcasm. For future research, we suggest integrating an irony detection module to enhance the model's accuracy and robustness in these complex cases.

4.2 Predominant Polarity

The first key aspect in the analysis of the expected results is the predominant polarity of the selected songs (Figure 1). A differentiated distribution of the feelings and emotions evoked by each musical piece is expected, based on its lyrical and musical characteristics.

Zero: This song, characterized by introspective and possibly more somber lyrics, will likely evoke a greater proportion of negative feelings, such as disdain, frustration, or anguish. The key, tempo, and lyrical content contribute to a more introspective and perhaps somewhat melancholic atmosphere. This type of lyrics, which tend to explore vulnerability or internal pain, generally generates a more negative polarity in the emotional responses of listeners. For Generation Z, whose relationship with music can be influenced by a mix of feelings of anxiety and disconnection, songs of this type resonate by offering an outlet or identification with similar personal experiences.

Dafne: This song is anticipated to generate a more neutral response in terms of emotional polarity. Although its tone may be tinged with melancholy or sadness, the lyrics and musical style could be designed in a way that does not excessively delve into pain, but rather invites reflection. This could result in an ambivalent feeling in listeners, who experience a mixture of nostalgia and serenity, but without as negative an emotional charge as "Zero" might offer. The use of a minor key and a moderate tempo could reinforce that feeling of gentle introspection, which makes the song interpreted by many as thoughtful,

but not completely depressing.

Al Voluntario: The collaborative song with Paty Cantú, due to its more accessible and positive style, is expected to generate a positive emotional polarity, favoring feelings of joy, optimism, or hope. Musical collaborations, especially with artists known for their cheerful and energetic pop style, usually produce a higher emotional response, especially when the music reflects a message of overcoming, love, or community. In addition, the collaboration itself is an element that could bring a feeling of union and solidarity between the artists and their listeners, which particularly resonates with Generation Z, which deeply values connection and collectivity.

These differences in the expected emotional polarity reflect how the musical elements (tempo, key, dynamics) and the song lyrics contribute to a completely different auditory experience, even when the songs may share superficial characteristics such as genre or format (Figure 4).

4.3 Emotional Profile: Differences in Primary Emotions

The analysis of the emotional profile involves understanding the primary emotions that each song evokes in listeners, differentiating between negative and positive emotions, and how these are associated with both the lyrics and the musical attributes, as shown in Figure 2.

Zero: The introspection and somber tone of the lyrics of "Zero" could primarily evoke negative emotions, especially sadness or disdain. Generation Z, which lives in a context marked by a high emotional load (between social, political, and economic uncertainty), often seeks in music a way to process and make sense of those complex emotions. Therefore, "Zero" could connect with listeners who experience internal frustration or emotional disconnection, creating a predominantly sad emotional profile, but also with a nuance of disdain or anger that can be more subtle but equally significant.

Dafne: The song "Dafne", which presents a more moderate and reflective tone, will probably generate a combination of emotions such as sadness and nostalgia. Unlike "Zero", which could trigger a stronger feeling of anguish, "Dafne" invites a gentle reflection, so the associated emotions are more melancholic than purely negative. However, this combination of light sadness with moments of calm can make listeners experience emotions of serenity or reflection, without reaching

the emotional intensity of a song like "Zero".

Collaboration with Paty Cantú: The collaboration with Paty Cantú can lead to a very different emotional profile, with predominantly positive emotions, such as joy, surprise, and anticipation. The presence of an artist known for their positive energy and accessible style can generate a feeling of happiness and motivation. It is likely that the combination of voices, the energetic dynamics of the song, and its possible message of hope or community will evoke more optimistic and unifying feelings, key elements for Generation Z, which tends to value connection and collaboration.

These differences in the emotional profile do not only depend on the lyrics and the theme of the song but also on the interaction with musical characteristics such as tempo, key, and dynamics, which reinforce the expected primary emotions.

4.4 Implications: Insights for Music Producers

The analysis of emotional patterns and Generation Z's responses to different types of music and lyrics has significant implications for music producers. First, it is important to recognize that Generation Z, more than any other previous generation, is deeply connected to music not only as a means of entertainment but as a form of emotional and personal expression. Songs that address themes such as vulnerability, self-exploration, or personal overcoming have a particular resonance with this audience.

For music producers, the insights derived from this research can guide decisions regarding the creation and production of future songs, especially those that involve collaborations or introspective themes. Songs that seek to capture a positive emotional response should focus on elements such as energetic rhythms, major keys, and messages of solidarity or overcoming. On the other hand, those songs that seek to connect with feelings of sadness or reflection, as in the case of "Zero" or "Dafne", can benefit from slower tempos, minor keys, and lyrics that speak of vulnerability or contemplation.

Furthermore, musical collaborations seem to be a powerful tool for producers who seek to capture the attention of Generation Z, as these provide a feeling of union and community, something that particularly resonates in a context where social media and online participation are fundamental to the identity of this generation. Therefore, understanding how different duet styles affect emotional perception is crucial for producers who want to create songs that truly connect

with the young audience. In conclusion, this analysis provides a clear vision of how songs are not only perceived by their sound content but also by the emotions they evoke. A deep understanding of these dynamics allows music producers to create auditory experiences that not only accompany but also reflect and amplify the emotions of Generation Z.

5 Conclusions and Future Work

This work proposes a robust methodological framework to comprehensively analyze the sentiment that Generation Z manifests towards José Madero's musical collaborations, with a special emphasis on his recent duets. By combining advanced natural language processing techniques—such as the BERT model, which allows for capturing contextual nuances in the lyrics with the structured use of the emotional NRC lexicon—a dual approach is achieved: on the one hand, quantitative and computational; on the other, qualitative and musically contextualized. This methodological convergence is essential for understanding how emotions are articulated in cultural products aimed at a generation with high levels of digital literacy, immediate aesthetic sensitivity, and complex patterns of emotional identification.

The expected findings not only contribute to the academic understanding of the link between music and emotion but also offer practical tools for the music industry. By identifying how certain lyrical structures, tempo decisions, or specific collaborations alter the emotional perception of a song, new possibilities open up to design more precise artistic and marketing strategies, focused on the type of emotional resonance that this cohort demands. In particular, it is shown that emotional polarity and the intensity of certain emotions such as nostalgia or surprise play a decisive role in the popularity of a piece among young listeners, which can be capitalized on in future productions, tours, or digital dissemination campaigns.

Likewise, this study allows for the visualization of the role of shared emotionality in the construction of community among fans, a key phenomenon in the era of social media. Music no longer operates in a vacuum; it is commented on, interpreted, appropriated, and even transformed into memetic or narrative content within spaces like TikTok, Instagram, and X (formerly Twitter). Understanding what type of duets generate greater emotional engagement can, therefore, facilitate the creation of more immersive, transmedia, and viralizable musical experiences.

5.1 Limitations and Suggestions for Future Work

We acknowledge the potential for biases in both the dataset and model predictions. The primary source of bias in the dataset stems from its narrow collection scope: comments were sourced from a single platform and relate to a specific artist and his collaborations. This may not represent the full spectrum of opinions from the Generation Z demographic, which is culturally and geographically diverse. Furthermore, the selection of comments from a specific private university introduces a demographic and socioeconomic bias, limiting the generalizability of the findings.

To mitigate these issues in future work, the following strategies are suggested:

- **Expand the Dataset:** Future research should aim to collect data from a wider range of platforms (e.g., TikTok, Reddit) and diverse geographical regions to capture a more representative sample of the target demographic.
- **Address Algorithmic Bias:** While the PSO-BERT model performed well, it may still exhibit inherent biases related to its training data. A detailed error analysis of false positives and false negatives could help identify specific linguistic patterns or cultural nuances that the model struggles with.
- **Incorporate Ethical AI Frameworks:** Integrating ethical AI frameworks throughout the research process—from data collection to model deployment—can help proactively identify and reduce biases. This includes regular audits for fairness and the development of more transparent, explainable models.

5.2 Future Work

From this framework, several lines of future work are envisioned. First, it is proposed to expand the analysis base to other duets by contemporary Latin American artists, to identify common or divergent emotional patterns according to musical genre, nationality, or type of collaboration. Second, it is suggested to incorporate visual analysis tools (such as the detection of facial expressions or real-time reactions on video platforms) to complement the textual reading of sentiment with an audiovisual dimension. This would allow for the construction of even more complete affective maps, integrating multisensory data.

As it was an exploratory study, this approach enabled us to mitigate external variables and focus on the

gender and sentiment dynamics within a group with similar demographic and cultural characteristics, providing a solid foundation for future research that seeks broader generalization. A broader sample is interesting to investigate how this can be reflected in different communities, cultures, regions, races or religious background.

Furthermore, it would be relevant to explore how the emotions induced by these duets vary depending on demographic variables such as gender, sexual orientation, socioeconomic level, or geographical context, which would open a necessary dialogue between music, identity, and cultural belonging. Finally, it is recommended to develop personalized music recommendation interfaces that integrate these findings, aimed at providing Generation Z listeners with an emotionally coherent and enriched experience. In summary, this study constitutes an initial step towards a deep understanding of musical emotions in a digitized context and offers an open platform for interdisciplinary research that combines computational linguistics, musical analysis, affective psychology, and cultural studies. Music, as an emotional and social vehicle, remains fertile ground for innovating the way we understand and connect with each other.

Data Availability Statement

List of source songs:

1. Zero (<https://www.youtube.com/watch?v=1pVbRBuP7tl>)
2. Dafne (<https://www.youtube.com/watch?v=ZkQW1mF0eBM>)
3. Al voluntario (<https://www.youtube.com/watch?v=YxgJ6lqjVxl>)
4. - es + (Menos es Más) (<https://www.youtube.com/watch?v=U9IRqmFJvZw>)
5. Codependientes (<https://www.youtube.com/watch?v=K4bV6qgk2zs>)
6. Te soñé (https://www.youtube.com/watch?v=7sG_JXcJp8Q)
7. Rómpase el Vidrio en Caso de Emergencia (<https://www.youtube.com/watch?v=3I4yQNfX2h8>)

1

¹The videos may be removed from YouTube without previous notice.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Makhdom, N., Verma, H. N., & Yadav, A. K. (2024). A review on sentiment and emotion analysis for computational literary studies. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*. [CrossRef]
- [2] Shen, X. (2024). Sentiment analysis of modern Chinese literature based on deep learning. *JES*. [CrossRef]
- [3] Ghatiya, L., Dhaketa, M., Saraf, K., Neema, J., & Vatsa, M. (2024). MoodSphere an approach to analyze sentiments using deep learning. *International Research Journal of Modernization in Engineering Technology and Science*. [CrossRef]
- [4] Horvat, M., Gledec, G., & Leontić, F. (2024). Hybrid natural language processing model for sentiment analysis during natural crisis. *Electronics*, 13(10), 1991. [CrossRef]
- [5] Ravichandiran, S. (2021). *Getting started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*. Packt Publishing Ltd.
- [6] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135. [CrossRef]
- [7] Maas, A., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 142-150).
- [8] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1631-1642).
- [9] Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.
- [10] Samih, A., Ghadi, A., & Fennan, A. (2023). Enhanced sentiment analysis based on improved word embeddings and XGboost. *International Journal of Electrical and Computer Engineering (IJECE)*, 13(2), 1827-1836. [CrossRef]
- [11] Ferdous, S. M., Shah Newaz, S. H., Shabbir Mugdha, S. B., & Uddin, M. (2024). Sentiment analysis

- in the transformative era of machine learning: A comprehensive review. *Statistics Optimization & Information Computing*. [CrossRef]
- [12] Wu, Y., Jin, Z., Shi, C., Liang, P., & Zhan, T. (2024). Research on the application of deep learning-based BERT model in sentiment analysis. *Applied and Computational Engineering*. [CrossRef]
- [13] Chaisen, T., Charoenkwan, P., Kim, C. G., & Thiengburanathum, P. (2024). A zero-shot interpretable framework for sentiment polarity extraction. *IEEE Access*. [CrossRef]
- [14] Moudhich, I., & Fennan, A. (2024). Graph embedding approach to analyze sentiments on cryptocurrency. *International Journal of Electrical and Computer Engineering (IJECE)*, 14(1), 690-697. [CrossRef]
- [15] Dhole, Y. B., Sharma, D. K., Patil, S. B., & Chavan, D. (2024). Unveiling market sentiments: Finbert-powered analysis of stock news headlines. *International Journal of Scientific Research in Engineering and Management*. [CrossRef]
- [16] Ilmawan, L. B., Muladi, M., & Prasetya, D. D. (2024). Negation handling for sentiment analysis task: Approaches and performance analysis. *International Journal of Electrical and Computer Engineering (IJECE)*, 14(3), 3382-3393. [CrossRef]
- [17] Liu, B. (2010). *Sentiment analysis and subjectivity*. Synthesis Lectures on Human Language Technologies. [CrossRef]
- [18] Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82-89.
- [19] Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 48-57.
- [20] Zeng, Y., Li, Z., Chen, Z., & Ma, H. (2023). Aspect-level sentiment analysis based on semantic heterogeneous graph convolutional network. *Frontiers of Computer Science*, 17(6), 176340.
- [21] Chu, M., Chen, Y., Yang, L., & Wang, J. (2022). Language interpretation in travel guidance platform: Text mining and sentiment analysis of TripAdvisor reviews. *Frontiers in Psychology*. [CrossRef]
- [22] Gajbhiye, P., & Agrawal, A. J. (2022). Pragmatic approach for Twitter analysis and perform prediction. *International Journal of Health Sciences*. [CrossRef]
- [23] Joseph, T. (2024). Natural language processing (NLP) for sentiment analysis in social media. *International Journal of Computing and Engineering*. [CrossRef]
- [24] Yacoub, A. D., Slim, S., & Aboutabl, A. (2024). A survey of sentiment analysis and sarcasm detection: Challenges, techniques, and trends. *International Journal of Electrical and Computer Engineering Systems*, 15(1), 69-78.
- [25] Bazai, S. U., Marjan, S., Aslam, S., Amphawan, A., & Neo, T. K. (2023). A comprehensive survey on sentiment analysis techniques. *International Journal of Technology*, 14(6). [CrossRef]
- [26] Low, H. Q., Keikhosrokiani, P., & Asl, M. P. (2024). Decoding violence against women: Analysing harassment in Middle Eastern literature with machine learning and sentiment analysis. *Humanities and Social Sciences Communications*. [CrossRef]
- [27] Khalid, E. T., Khalefa, M. S., Yassen, W., & Yassin, A. A. (2023). Omicron virus emotions understanding system based on deep learning architecture. *Journal of Ambient Intelligence and Humanized Computing*. [CrossRef]
- [28] Mermer, G., & Özsezer, G. (2022). Discussions about COVID-19 vaccination on Twitter in Turkey: Sentiment analysis. *Disaster Medicine and Public Health Preparedness*. [CrossRef]
- [29] He, W., Yuan, Q., & Li, N. (2023). Research on the characteristics of internet public opinion and public sentiment after the Sichuan earthquake based on the perspective of Weibo. *Applied Sciences*, 13(3), 1335. [CrossRef]
- [30] Chen, M., Ubul, K., Xu, X., Aysa, A., & Muhammad, M. (2022). Connecting text classification with image classification: A new preprocessing method for implicit sentiment text classification. *Sensors*, 22(5), 1899. [CrossRef]
- [31] Jin, C., Liu, R., Tang, B., & Cai, B. (2023). Predict FTSE100 stock movements using business news sentiment and machine learning. *Theoretical and Natural Science*. [CrossRef]
- [32] Sultan, S., Javaid, Q., Rehman, E., Alahmadi, A. A., Ullah, N., & Khan, W. (2022). MELex: The construction of Malay-English sentiment lexicon. *Computers Materials & Continua*. [CrossRef]
- [33] Alshahrani, H. J., Hassan, A. Q. A., Almalki, N. S., Alnfai, M. M., Salama, A. S., & Hamza, M. A. (2023). Applied linguistics with red-tailed hawk optimizer-based ensemble learning strategy in natural language processing. *IEEE Access*. [CrossRef]
- [34] Mohammad, S. M., & Turney, P. D. (2013). *NRC Emotion Lexicon*. National Research Council Canada. [CrossRef]
- [35] Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1, pp. 216-225).
- [36] Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation* (pp. 2200-2204).
- [37] Eberhart, R., & Kennedy, J. (1995). Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks* (Vol. 4, pp. 1942-1948).



Dr. Alberto Ochoa-Zezzatti He is a Full Professor and Researcher at the Universidad Autónoma de Ciudad Juárez (UACJ), recognized with SNI Level 2 and PRODEP distinction. He holds a PhD and has completed multiple postdoctoral stays in areas such as energy, industry, and artificial intelligence. He has authored over 575 scientific publications in seven languages, with more than 2,100 citations, and has supervised 97 PhD, 92 master's, and 57 undergraduate theses. He is an active member of the Academic Board (NAB) of UACJ's Technology PhD Program (PNPC). His research interests include applied artificial intelligence, data science, evolutionary computing, natural language processing, and social modeling for smart cities and Industry 4.0. He has led national and international projects, including a European Union initiative on road safety and a collaborative AI project with Uruguay, as well as local efforts like wildfire prevention in Mexico. He has conducted research stays at COMIMSA, ITAM, the University of Zaragoza, and is currently on sabbatical at Universidad Anáhuac, contributing to the Master's in Artificial Intelligence. He has organized the HIS Workshop at MICAI since 2008, serves as a reviewer for journals such as Applied Soft Computing and Computers in Human Behavior, and authored a book published by Taylor & Francis. (Email: alberto.ochoa@uacj.mx)



Dr. Huberto Garcia Castellanos was born in Ignacio Allende Village, Durango, Mexico in 1960. Currently he lives in El Paso, Texas, USA. He received the B.S Instituto Tecnológico de Durango in Durango, Mexico. The M.S and PhD degrees Industrial Engineering from the Tecnológico Nacional de México/Instituto Tecnológico de Ciudad Juárez Mexico. From 1991 to date, he has been a professor in Tecnológico Nacional de México/ITCJ. He has many publications in the United States of America, England, Mexico, Egypt, Morocco, Switzerland, and India. He has had professional experience in several automotive international manufacturers' positions like General Motors, Nissan, Ford, and Chrysler for fourteen years. He developed many functions in several basic levels, production supervisor, industrial Engineer, Product Engineer, Quality Engineer, Production Manager, Environment Engineer Industries Auditor. He participates as a researcher for a location plant for residuals urban water in Ciudad Juárez, Mexico. Also, their publications are about industrial projects management success, artificial intelligence and logistics chain supply, and other automotive topics publications. He is a reviewer for the Dyna journal in Europe.



Dr. Roberto Contreras-Masse Received his PhD in Technology from Universidad Autónoma de Ciudad Juárez (UACJ). He is a research professor at TecNM Campus ITCJ and a member of Mexico's National System of Researchers. With over 25 years of experience, Roberto teaches Artificial Intelligence and has improved year after year the syllabus to match current changes of the field. His expertise spans a wide range of topics within technology, and has special interest on artificial intelligence, computer vision, signal processing and intelligent data analysis. He serves as reviewer for journals from Springer and Elsevier. (Email: roberto.cm@cdjuarez.tecnm.mx)

Roberto Contreras-Moheno is currently an undergraduate student with academic interests in medicine, biomedical engineering, and music. He graduated with honors from Instituto México High School, where he was awarded distinction cords in both STEM and the Arts (for his achievements in music). During high school, he also served as lead pianist, highlighting his commitment to artistic excellence. Accepted into the University of Texas at El Paso (UTEP) for academic excellence, he is now beginning to explore scientific research and is committed to developing a strong foundation in interdisciplinary inquiry.