

Análisis de sentimientos expresados en la música producida en Colombia entre 1990 a 2023 como reflejo de la coyuntura social y política de la época.

Sentiment Analysis of Emotions Expressed in Music Produced in Colombia between 1990 and 2023 as a Reflection of the Social and Political Context of the Period.

Leidy Johanna Bejarano Rodríguez,¹ Zulmira Mayerly Santos Pérez,¹ Wilmar Leonardo Álvarez Guzmán¹

¹ Universidad Central de Colombia

{lbejaranor4, zsantosp, walvarezg1}@ucentral.edu.co

Resumen: Colombia ha sido un país históricamente marcado por el conflicto armado, cuyas dinámicas han influido en distintas formas de expresión cultural, entre ellas la música. Este estudio explora la relación entre el contenido emocional de las letras de canciones producidas en Colombia entre 1990 y 2023 y la violencia asociada al conflicto armado, medida mediante el número anual de víctimas de masacres. Para ello, se construyó un corpus de 3.400 canciones, analizado mediante técnicas de Procesamiento de Lenguaje Natural (PLN) y clasificación automática con el Modelo de Lenguaje de Gran Tamaño (LLM), DeepSeek R1. Los resultados muestran un predominio de polaridades negativas y emociones asociadas principalmente con tristeza y alegría, así como una recurrencia de sentimientos relacionados con la melancolía. Adicionalmente, el análisis de correlación de Spearman identificó asociaciones monotónicas moderadas y estadísticamente significativas entre el número de víctimas y algunas categorías emocionales. No obstante, debido a las limitaciones inherentes a la clasificación automática y al carácter exploratorio del estudio, estos hallazgos deben interpretarse como aproximaciones exploratorias y no como relaciones concluyentes. En este sentido, el trabajo propone una primera aproximación interdisciplinaria al análisis computacional de expresiones emocionales en la música colombiana y resalta el potencial del PLN para el estudio de fenómenos culturales e históricos complejos.

Palabras clave: Análisis de sentimientos, Modelo de Lenguaje de Gran Tamaño, Letras de canciones, Procesamiento de Lenguaje Natural.

Abstract: Colombia has historically been marked by armed conflict, whose dynamics have influenced various forms of cultural expression, including music. This study explores the relationship between the emotional content of song lyrics produced in Colombia between 1990 and 2023 and conflict-related violence, measured through the annual number of massacre victims. To this end, a corpus of 3,400 songs was analyzed using Natural Language Processing (NLP) techniques and automatic classification with the Large Language Model (LLM) DeepSeek R1. The results show a predominance of negative polarities and emotions mainly associated with sadness and joy, as well as a recurrent presence of sentiments related to melancholy. In addition, Spearman correlation analysis identified moderate and statistically significant monotonic associations between the number of victims and certain emotional categories. However, due to the limitations inherent in automatic classification and the exploratory nature of the study, these findings should be interpreted as exploratory approximations rather than conclusive relationships. In this sense, the study provides a first interdisciplinary approach to the computational analysis of emotional expressions in Colombian music and highlights the potential of NLP for the study of complex cultural and historical phenomena.

Keywords: Sentiment analysis, Large Language Models, Song lyrics, Natural Language Processing.

1 *Introducción*¹

La música ha sido un medio de expresión cultural que ha permitido a las sociedades comunicar tradición e historia (Castro Rueda & Quishpe Gaibor, 2019), así como expresar emociones, inconformidades y anhelos, reflejando las dinámicas sociales y políticas de cada época. En este sentido, las letras musicales han funcionado como un espacio donde se manifiestan problemáticas como la desigualdad, la violencia, la resistencia y la esperanza, consolidándose como un medio clave para interpretar procesos históricos y sociales. En este sentido, comprender la música implica analizar la relación entre la realidad musical y las características de la sociedad que la crea o interpreta (Hormigos, 2010).

Sin embargo, la música suele analizarse desde enfoques predominantemente subjetivos, en los que las letras son interpretadas sin considerar plenamente el contexto social o histórico del que provienen. Esto puede conducir a la estigmatización de ciertos géneros o composiciones (Tagg, 2012) y limita la identificación sistemática de patrones emocionales. En este contexto, el uso de herramientas como el Procesamiento de Lenguaje Natural (PLN) y el análisis de sentimientos basado en Inteligencia Artificial (IA) permite superar estas limitaciones (Rabino, 2024), al posibilitar el análisis automatizado de la carga emocional presente en textos como las letras de canciones.

En este tipo de enfoques, el análisis del contenido emocional permite identificar la polaridad de los textos, así como las emociones y sentimientos expresados en ellos. La polaridad se entiende como la orientación valorativa del contenido en términos positivos, negativos o neutros, asociada con la intención y el tono emocional del autor (Velázquez Carrillo, 2021). Las emociones corresponden a respuestas automáticas y de corta duración ante estímulos (Álvarez, 2024), de acuerdo a Paul Ekman y Wallace Friesen, existen seis emociones básicas universales y son alegría, tristeza, enojo, miedo, sorpresa y asco (Leperski, 2017). Por su parte, los sentimientos representan estados afectivos más complejos y prolongados, derivados de procesos cognitivos que generan una experiencia emocional más

duradera (Álvarez, 2024). En este contexto, el análisis de sentimientos, como técnica del PLN, permite extraer información subjetiva de los textos con el fin de identificar su orientación emocional en distintos niveles de análisis – como documento, oración o aspecto –, facilitando el procesamiento sistemático de grandes volúmenes de información con mayor consistencia y menor sesgo interpretativo frente al análisis exclusivamente humano (Sobrino Sande, 2018).

A pesar de los avances en este campo, en el contexto colombiano los estudios que articulan el análisis computacional de letras musicales con el entorno social y político siguen siendo limitados. Investigaciones como la de Muñoz López (2021) evidencian el papel de la música en procesos de memoria y reparación en escenarios de violencia; sin embargo, estos enfoques han sido principalmente cualitativos y no permiten identificar patrones emocionales de manera sistemática en grandes volúmenes de datos. De manera similar, estudios sobre la resignificación de géneros musicales – como el heavy metal en Colombia – han evidenciado su uso como forma de denuncia y resistencia frente al conflicto armado (Becerra González, 2022), aunque sin incorporar herramientas computacionales para el análisis masivo de sus contenidos.

En este contexto, el presente estudio tiene como objetivo analizar la polaridad, emociones y sentimientos presentes en las letras de canciones producidas en Colombia entre 1990 y 2023, con el fin de explorar su posible relación con la coyuntura social y política del periodo. Para ello, se emplean técnicas de PLN y Modelos de Lenguaje de Gran Tamaño (LLM), utilizados como una aproximación exploratoria para identificar posibles tendencias emocionales en un corpus de canciones. De esta manera, el estudio busca aportar una perspectiva interdisciplinaria al análisis de la música como expresión cultural vinculada a contextos históricos y sociales específicos.

2 *Análisis de sentimientos en expresiones culturales y musicales*

El análisis de sentimientos ha sido ampliamente aplicado en diversos contextos de investigación, especialmente en el estudio de opiniones y percepciones de usuarios en entornos digitales, así como en el análisis de discursos políticos y redes sociales. Sin embargo, su aplicación en

¹ Este artículo deriva de un trabajo de grado de maestría.

productos culturales, particularmente en la música, ha sido más limitada, especialmente en lo que respecta a su articulación con el contexto histórico, social y cultural.

Entre estos estudios se destaca el de Palomeque y De Lucio (2021), quienes analizaron la evolución del sentimiento expresado en las letras de canciones incluidas en el Billboard Hot 100 entre 1950 y 2020. Este trabajo también buscó determinar si el sentimiento expresado en la letra era independiente o se ajustaba a la música que la acompañaba. Los resultados mostraron una correlación limitada entre la valencia emocional de la música y la de las letras, así como una disminución del contenido lírico positivo a lo largo del tiempo.

De manera similar, García Vázquez et al. (2023) aplicaron modelos de aprendizaje profundo para la clasificación automática de sentimientos en letras de canciones en español, evidenciando la viabilidad de las redes neuronales para este tipo de tareas. No obstante, los autores señalan que el uso de modelos más robustos y mecanismos de atención podría mejorar los resultados, especialmente en el manejo de secuencias textuales más largas.

En el caso de Srinilta et al. (2017), quienes analizaron letras de canciones en idioma tailandés, evidenciaron tanto el potencial como las limitaciones del análisis de sentimientos en lenguas con baja disponibilidad de recursos lingüísticos. En este contexto, el desempeño de los enfoques utilizados se ve condicionado tanto por las características del idioma como por el tipo de modelos empleados, lo que restringe su capacidad para capturar adecuadamente las expresiones emocionales.

Mientras que algunos estudios evidencian limitaciones en el análisis de sentimientos en lenguas con escasos recursos lingüísticos, otros muestran resultados más favorables cuando se desarrollan enfoques adaptados al contexto del idioma. En este sentido, Albán Morales y Gualoto Fuentes (2024) desarrollaron un algoritmo para el análisis de sentimientos en idioma Kichwa basado en la construcción de un corpus propio y en el uso de técnicas de similitud textual, alcanzando altos niveles de precisión en la clasificación. Este contraste sugiere que el desempeño de los modelos no depende únicamente de la disponibilidad de recursos lingüísticos, sino también del diseño metodológico y de la adecuación de las herramientas al contexto específico del idioma.

Esta situación también ha sido evidenciada por Susatama Ruiz et al. (2022), quienes, aunque no trabajaron con letras de canciones sino con publicaciones en Twitter, identificaron dificultades para interpretar expresiones coloquiales propias del contexto colombiano. Estas limitaciones afectaron la precisión de los modelos y pusieron de manifiesto la necesidad de enfoques adaptados al idioma y al contexto cultural.

Otros enfoques han abordado la emoción musical desde el procesamiento directo de la señal sonora. Han et al. (2023) desarrollaron modelos de aprendizaje profundo para el reconocimiento automático de emociones musicales a partir de características acústicas, combinando distintas arquitecturas neuronales para capturar tanto patrones del sonido como su evolución temporal. Sus resultados evidencian mejoras frente a modelos tradicionales en la identificación automática de emociones musicales.

En complemento a estos estudios empíricos, Patil et al. (2025) presentan una revisión exhaustiva de los métodos, tecnologías y conjuntos de datos empleados en el análisis de emociones musicales, destacando los avances en modelos profundos y enfoques multimodales. No obstante, señalan que aún persisten desafíos relacionados con la subjetividad de la emoción y la influencia del contexto cultural.

En conjunto, estos estudios evidencian avances significativos en el uso de técnicas de PLN para el análisis de emociones en música y otros contextos, así como limitaciones asociadas al idioma, la disponibilidad de recursos lingüísticos y la influencia del contexto cultural. Sin embargo, persiste un vacío en investigaciones que integren de manera sistemática el análisis de sentimientos en letras de canciones con su contexto histórico y sociopolítico, particularmente en el caso colombiano. En este sentido, resulta pertinente desarrollar estudios que permitan comprender la relación entre la expresión emocional en la música y las dinámicas sociales en las que esta se produce.

3 Metodología

3.1 Enfoque metodológico

El estudio se desarrolló bajo el enfoque CRISP-DM, el cual permitió estructurar el proceso de análisis en fases de comprensión del problema,

preparación de datos, modelado y evaluación. Dado que el objetivo no fue el despliegue de un sistema operativo, esta fase no fue considerada.

3.2 Fuentes de datos

Para la construcción del corpus de letras de canciones se exploraron inicialmente alternativas automatizadas de recolección de información mediante APIs de plataformas digitales y búsquedas web; no obstante, estas opciones presentaron limitaciones metodológicas importantes. La mayoría de estas herramientas requerían información previa sobre artistas, canciones o fechas de publicación, lo que podía introducir sesgos en la selección del corpus al favorecer canciones más populares o con mayor presencia digital. Adicionalmente, la disponibilidad de información para las décadas de 1990 y comienzos de los 2000 era considerablemente más limitada, dificultando la construcción de un corpus históricamente equilibrado y representativo para todo el periodo analizado.

En consecuencia, se optó por utilizar la base de datos suministrada por la Dirección Nacional de Derecho de Autor (DNDA),² compuesta por 409.473 registros y cuatro variables: título de la canción, año de registro, autor y ciudad. Tras un proceso de limpieza y depuración, la base quedó conformada por 383.341 registros válidos.

A partir de este conjunto se seleccionaron cien canciones por año, construyendo un corpus final de 3.400 registros para el periodo 1990-2023. La selección respondió a un criterio operativo basado en la disponibilidad de enlaces en la plataforma YouTube, necesarios para la posterior extracción de audio y transcripción.

De manera complementaria, se incorporó la base de datos de casos de masacres en Colombia del Observatorio de Memoria y Conflicto del Centro Nacional de Memoria Histórica (CNMH), con 4.627 registros para el periodo 1947-2024, con el fin de contextualizar los resultados del análisis en relación con eventos sociopolíticos.

² La Dirección Nacional de Derecho de Autor es un organismo del Estado Colombiano que se encarga del diseño, dirección, administración y ejecución de las políticas gubernamentales en materia de derecho de autor y derechos conexos (Dirección Nacional de Derecho de Autor [DNDA], 2024).

3.3 Construcción y preparación del corpus

La construcción del corpus implicó un proceso de limpieza, depuración y estandarización de la base de datos de la DNDA. En primer lugar, se desagregaron registros que contenían múltiples títulos de canciones en una misma fila, asociados a un mismo autor y año, con el fin de garantizar que cada registro correspondiera a una única canción. Posteriormente, se filtraron los datos correspondientes al periodo 1990-2023 y se eliminaron registros con inconsistencias en la variable temporal.

A partir de la base depurada, se intentó obtener de manera automatizada las URL de las canciones en YouTube mediante su API, realizando consultas a partir del título de la canción y el nombre del compositor. Sin embargo, este procedimiento no arrojó resultados suficientemente confiables debido a varias limitaciones. En primer lugar, la base de datos de la DNDA registra compositores, mientras que en plataformas como YouTube las canciones suelen estar asociadas al nombre del intérprete, dificultando la recuperación automática de coincidencias precisas. Adicionalmente, muchos títulos de canciones no son únicos, lo que generaba ambigüedades y resultados incorrectos al recuperar enlaces correspondientes a canciones distintas, pero con nombres similares o idénticos.

En consecuencia, se optó por realizar manualmente la búsqueda, validación de enlaces y clasificación de géneros musicales, garantizando una mayor correspondencia entre el registro original y la canción recuperada. Este procedimiento permitió asegurar la disponibilidad de al menos cien canciones por año para el periodo analizado, posibilitando posteriormente la extracción de audio y la transcripción del corpus.

Una vez obtenidos los enlaces, se descargaron los audios en formato .wav mediante la librería `yt_dlp`, los cuales fueron posteriormente transcritos utilizando herramientas de reconocimiento de voz. Para ello, se emplearon las librerías `speech_recognition` y `Vosk`, esta última utilizada principalmente debido a su capacidad de procesamiento *offline* y ausencia de restricciones en volumen de datos, lo que permitió cubrir la totalidad del corpus. En casos puntuales donde la transcripción automática no fue posible, las letras fueron completadas manualmente.

Finalmente, con el fin de facilitar el análisis, los géneros musicales fueron agrupados en cinco categorías generales: Folclor y Tradición Colombiana, Música Popular y Regional Latinoamericana, Música Comercial, Temática y Moderna, Rock y Música Urbana, a partir de criterios de similitud cultural y musical.

3.4 Estrategia de modelado

Para la clasificación de polaridad, emociones y sentimientos, se consideró inicialmente un enfoque de aprendizaje supervisado tradicional, mediante la implementación de modelos basados en Redes Neuronales Convolucionales (CNN) y redes recurrentes del tipo Long Short-Term Memory (LSTM). Para ello, se utilizó un conjunto de datos de tweets etiquetados, adecuado para tareas de clasificación textual en estas tres dimensiones.

No obstante, las características lingüísticas de este corpus –compuesto por textos breves, directos y de baja complejidad sintáctica– diferían sustancialmente de la estructura semántica y narrativa de las letras de canciones, lo que anticipaba limitaciones en la capacidad de generalización de los modelos hacia este dominio.

3.4.1 Modelos supervisados (CNN, LSTM)

Con el fin de evaluar empíricamente esta hipótesis, se entrenaron modelos CNN y LSTM utilizando el corpus de tweets como línea base. Los resultados evidenciaron un desempeño limitado en ambos casos: mientras que el modelo CNN presentó un marcado sobreajuste con una alta precisión en entrenamiento pero una caída significativa en validación; el modelo LSTM mostró un comportamiento de subajuste, sin lograr capturar patrones relevantes incluso en el propio dominio de los tweets.

En términos de desempeño, el modelo CNN alcanzó una precisión del 76,2 % en entrenamiento, pero descendió al 11,8 % en validación, evidenciando una baja capacidad de generalización. Por su parte, el modelo LSTM presentó bajos niveles de precisión tanto en entrenamiento (7,2 %) como en validación (5,6 %), lo que indica una limitada capacidad de aprendizaje.

En conjunto, estos resultados evidencian las limitaciones de los enfoques supervisados tradicionales incluso sobre textos breves y estructuralmente simples como los tweets utilizados durante el entrenamiento. En este

sentido, se consideró que la aplicación de estas arquitecturas a letras de canciones podría enfrentar dificultades adicionales, debido a la mayor complejidad semántica, presencia de metáforas, ambigüedad lingüística y carga simbólica característica del lenguaje musical, así como a la limitada disponibilidad de corpus etiquetados especializados en este dominio.

3.4.2 Modelos de lenguaje (LLM)

En consecuencia, se optó por emplear Modelos de Lenguaje de Gran Tamaño (LLM), debido a su capacidad de comprensión semántica y contextual, lo que permite abordar tareas de clasificación textual sin requerir entrenamiento supervisado directo sobre el corpus de estudio.

Para la evaluación de estos modelos, se construyó una base de referencia mediante la clasificación manual del 10 % del corpus consolidado. Esta muestra fue seleccionada de forma aleatoria por año y etiquetada por un profesional en psicología, siguiendo la teoría de las emociones básicas propuesta por Paul Ekman (2013), que contempla seis emociones fundamentales: alegría, tristeza, ira, miedo, sorpresa y asco. A partir de estas categorías, se definieron sentimientos específicos asociados a cada emoción, así como la polaridad correspondiente (positiva, negativa o neutra).

Esta base de referencia fue utilizada como conjunto de evaluación para comparar el desempeño de distintos modelos de lenguaje en la tarea de clasificación de polaridad, emociones y sentimientos. Se evaluaron modelos de la familia Gemma (4B, 12B y 27B parámetros) y DeepSeek R1 (8B, 14B y 32B parámetros).

Dado que estos modelos no fueron entrenados desde cero, su implementación se realizó mediante técnicas de prompt engineering. Para ello, se diseñaron instrucciones estructuradas que definían el rol del modelo, la descripción de la tarea, las categorías de clasificación y el formato de salida esperado. Adicionalmente, se empleó una estrategia de Few-Shot Prompting basada en ejemplos previamente definidos, con el fin de orientar el proceso de clasificación y mejorar la consistencia de las respuestas generadas.

El proceso de clasificación se realizó utilizando la letra completa de cada canción como unidad de análisis, evitando segmentaciones que pudieran afectar la interpretación del contexto semántico global.

Asimismo, se adoptó un enfoque de etiqueta única predominante, en el cual el modelo asigna la categoría que mejor representa el contenido general de la canción.

3.5 Evaluación del desempeño

El desempeño de los modelos de lenguaje se evaluó comparando las clasificaciones generadas por cada modelo con la base de referencia construida a partir del 10 % del corpus, etiquetada manualmente por un profesional en psicología. Para ello, se emplearon métricas estándar de evaluación en tareas de clasificación: precisión, exhaustividad (Recall), F1-Score y coeficiente Kappa de Cohen, complementadas con una medida de asertividad definida como el porcentaje de coincidencia entre las etiquetas generadas por el modelo y la clasificación realizada por el profesional en psicología. Mientras las tres primeras permiten evaluar la calidad de la clasificación de los modelos, el coeficiente Kappa mide el nivel de acuerdo entre las clasificaciones generadas y la referencia experta, controlando la posibilidad de coincidencias atribuibles al azar (Flores, É., & Soriano, J., 2019).

La evaluación se realizó para las tareas de polaridad, emociones y sentimientos, utilizando la letra completa de cada canción como unidad de análisis. En términos generales, los modelos de la familia DeepSeek R1 obtuvieron mejores resultados que los modelos Gemma en las tres dimensiones evaluadas. Este comportamiento fue especialmente evidente en la clasificación de polaridad, donde se registraron los mayores niveles de concordancia con la referencia experta. Los valores obtenidos para cada modelo se presentan en la tabla 4 del anexo 1.

Dentro de los modelos evaluados, DeepSeek R1 32B alcanzó los mejores resultados en polaridad y emociones, mientras que DeepSeek R1 14B presentó los valores más altos en la clasificación de sentimientos. La evaluación de DeepSeek R1 32B mostró un comportamiento diferencial según la tarea analizada. La polaridad obtuvo los indicadores más favorables (F1-Score = 0,7323; Kappa = 0,5325), reflejando un nivel de acuerdo moderado con la clasificación realizada por la especialista. Por su parte, las emociones alcanzaron valores intermedios (F1-Score = 0,5930; Kappa = 0,3651), mientras que los

sentimientos registraron los niveles más bajos (F1-Score = 0,1503; Kappa = 0,1131).

La disminución progresiva del desempeño de los modelos entre las tareas de polaridad, emociones y sentimientos puede asociarse al aumento en la complejidad de la tarea. Mientras la polaridad se limita a tres categorías y las emociones a seis, la clasificación de sentimientos contempló cincuenta categorías distintas, muchas de ellas con significados próximos o parcialmente superpuestos. Esta cercanía semántica dificulta la diferenciación precisa entre clases y aumenta la incertidumbre en la asignación automática de etiquetas.

En términos cualitativos, se observó un mejor desempeño en canciones con lenguaje explícito y directo, mientras que la precisión disminuyó en contextos donde predominan recursos como metáforas, regionalismos o ambigüedad semántica. Esta limitación resulta particularmente relevante en géneros como la música urbana y el folclor, donde el significado suele estar mediado por contextos culturales específicos.

En conjunto, los resultados muestran que los modelos de lenguaje de gran tamaño constituyen una herramienta útil para la exploración computacional de contenidos emocionales en letras musicales. No obstante, los niveles de desempeño observados sugieren que los resultados deben interpretarse como aproximaciones analíticas sujetas a las limitaciones inherentes tanto al proceso de clasificación automática como a la complejidad semántica del dominio estudiado.

4 Resultados

4.1 Análisis de las masacres ocurridas en Colombia entre 1990 a 2023

Entre 1990 y 2023 se registraron en Colombia 3.483 masacres con un total de 19.455 víctimas. La evolución temporal (ver figura 1) evidencia una tendencia creciente durante la década de 1990, con un pico a comienzos de los años 2000, seguida de un descenso sostenido hasta 2019 y un posterior repunte en 2020.

Este comportamiento se asocia, de manera general, con fases de intensificación del conflicto armado interno durante la década de 1990 e inicios de los años 2000, así como con transformaciones posteriores derivadas de procesos de desmovilización y acuerdos de paz, y disputas recientes por el control territorial.

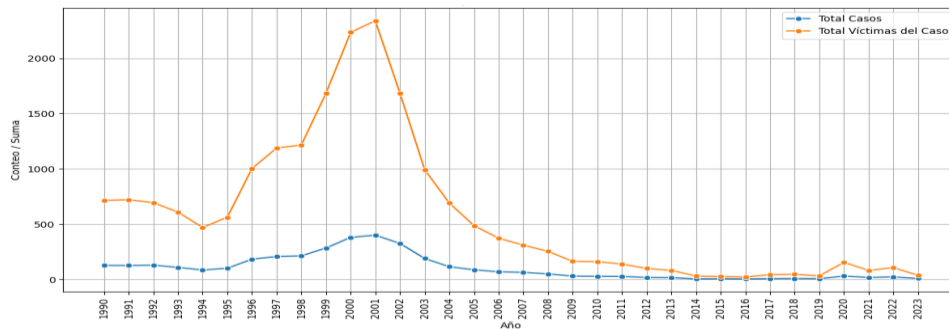


Figura 1: Serie anual de casos y víctimas de masacres en Colombia (1990-2023).

4.2 Resultados por género musical

El folclor y tradición colombiana concentra el mayor número de registros (ver tabla 5, anexo 2). De manera general, los géneros tradicionales presentan un contenido predominantemente afectivo y relacional, mientras que el rock y la música urbana muestran una mayor presencia de elementos asociados al conflicto, la denuncia y las experiencias cotidianas en contextos urbanos. Estas diferencias se examinan con mayor detalle en el análisis de polaridades, emociones y sentimientos.

4.3 Análisis emocional de las letras de las canciones

A continuación, se presentan los resultados de la clasificación realizada mediante el modelo DeepSeek R1 para las categorías de polaridad, emociones y sentimientos. Dado que se trata de una clasificación automática basada en modelos de lenguaje, las etiquetas asignadas deben interpretarse como aproximaciones al contenido emocional de las letras y no como representaciones exactas de su significado. Asimismo, la incertidumbre asociada a la clasificación aumenta con la complejidad de la tarea y el número de categorías consideradas, siendo menor en polaridad y mayor en sentimientos. En consecuencia, el análisis se aborda desde una perspectiva descriptiva y exploratoria, orientada a identificar tendencias emocionales y su evolución temporal dentro del corpus.

Adicionalmente, se examina la relación entre estas categorías y el número de víctimas de masacres mediante el coeficiente de correlación de Spearman, seleccionado debido a la no normalidad de las variables y a su capacidad para identificar asociaciones monotónicas entre series temporales; en

consecuencia, los resultados se interpretan como asociaciones estadísticas y no como evidencia de relaciones causales.

4.3.1 Polaridad

La clasificación de polaridad, entendida como la orientación predominante del mensaje en términos positivos, negativos o neutros, muestra un predominio de la polaridad negativa con 1.397 canciones (41,7 %), frente a las categorías positiva (30,2 %) y neutra (29,6 %).

Como se observa en la figura 2, la polaridad negativa mantiene una presencia constante a lo largo del periodo analizado. Aunque en algunos años iniciales (1990 y 1993) la polaridad positiva alcanza valores elevados, esta tendencia es posteriormente superada por la negativa, que presenta picos hacia finales de la década de 1990 e inicios de los 2000. Algunos de estos incrementos se presentan en años que también registraron elevados niveles de victimización (ver figura 1). A partir de 2010, la polaridad negativa continúa siendo predominante, aunque con una mayor alternancia frente a las categorías neutra y positiva.



Figura 2: Polaridad por año en relación del total de víctimas.

No obstante, la correlación de Spearman entre la polaridad y el número de víctimas muestra asociaciones débiles y estadísticamente no significativas para las tres categorías analizadas lo que indica que no existe evidencia suficiente para afirmar una relación sistemática entre ambas variables (ver tabla 1).

Polaridad	Correlación de Spearman	
	Coef	p-value
Negativo	0,19	0,28
Neutro	-0,31	0,08
Positivo	0,20	0,26

Tabla 1: Correlación de la polaridad y total de víctimas de masacres.

Estos resultados sugieren que la prevalencia de la polaridad negativa no puede atribuirse exclusivamente al contexto sociopolítico. La música incorpora de forma recurrente temáticas asociadas al desamor, la pérdida, la nostalgia y otros aspectos de la experiencia individual, por lo que las tendencias observadas podrían reflejar simultáneamente experiencias personales y elementos vinculados a contextos sociales más amplios, sin implicar relaciones causales directas.

El análisis por género muestra diferencias en la distribución de las polaridades: el folclor presenta una mayor proporción de valores positivos, el rock concentra la mayor carga negativa, el género urbano destaca por una mayor presencia de categorías neutras y la música comercial exhibe una distribución más equilibrada. Asimismo, las similitudes observadas en el vocabulario sugieren que las diferencias entre polaridades dependen no solo de las palabras utilizadas, sino también del contexto emocional en el que estas adquieren significado.

4.3.2 Emociones

Los resultados de la clasificación por emociones muestran un predominio de la tristeza a lo largo del periodo analizado (1.493 letras), consistente con la tendencia observada previamente para la polaridad negativa (ver figura 3). Esta emoción presenta incrementos hacia mediados de la década de 1990 y durante los primeros años del siglo XXI, periodos que también registraron elevados niveles de victimización. La alegría (1.056 registros) constituye la segunda emoción más frecuente y

mantiene una presencia relativamente constante a lo largo del periodo estudiado.

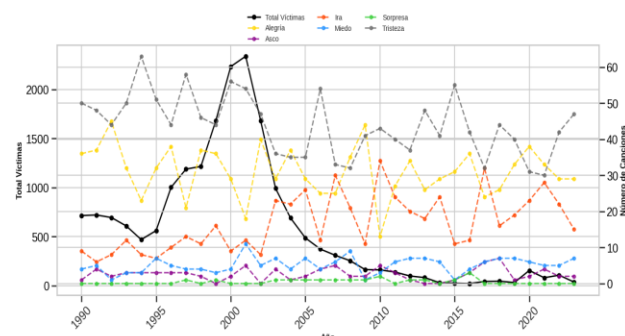


Figura 3: Emociones por año en relación del total de víctimas.

La ira (572 letras) exhibe una dinámica diferente. Mientras el número de víctimas disminuye a partir de los primeros años de la década de 2000, la proporción de canciones asociadas a esta emoción aumenta desde 2002 y se mantiene, con fluctuaciones, durante buena parte del periodo posterior. Este comportamiento podría estar relacionado con transformaciones en las dinámicas del conflicto armado, en las que otras formas de violencia –como las violaciones a los derechos humanos– adquirieron mayor relevancia.

Por su parte, las emociones de miedo (174 letras) y asco (88 letras) presentan un repunte alrededor de 2001, mientras que la sorpresa (17 letras) registra una presencia marginal dentro del corpus.

La correlación de Spearman entre el número total de víctimas y las emociones (ver tabla 2) permite complementar estos comportamientos.

Emoción	Correlación de Spearman	
	Coef	p-value
Alegría	0,17	0,33
Asco	0,03	0,88
Ira	-0,45	0,01
Miedo	-0,18	0,30
Sorpresa	-0,12	0,49
Tristeza	0,38	0,03

Tabla 2: Correlación de emociones y total de víctimas de masacres.

La tristeza presenta una correlación positiva de magnitud moderada y estadísticamente significativa ($p = 0,03$), lo que indica que los años con mayores niveles de victimización tienden a coincidir con una mayor proporción de canciones asociadas a esta emoción. Por su

parte, la ira muestra una correlación negativa moderada y significativa ($p = 0,01$), sugiriendo una disminución relativa de esta emoción en contextos de mayor violencia.

En contraste, las emociones de alegría y asco presentan asociaciones positivas débiles y no significativas ($p = 0,33$ y $p = 0,88$, respectivamente), mientras que miedo y sorpresa muestran asociaciones negativas débiles sin significancia estadística ($p = 0,30$ y $p = 0,49$). En el caso del asco y la sorpresa, los resultados deben interpretarse con cautela debido a su baja frecuencia en el corpus, lo que limita la robustez estadística del análisis.

El análisis por género evidencia diferencias en la distribución de emociones. Mientras el folclor y la música comercial presentan un predominio de tristeza y alegría, la música popular y el rock muestran una mayor presencia de tristeza e ira, reforzando su carácter crítico, y el género urbano combina tristeza, ira y alegría, configurando un discurso que articula experiencias cotidianas y problemáticas sociales.

A nivel léxico, aunque términos como “vida” y “amor” aparecen de forma transversal en todos los géneros, se observan matices diferenciados: en los géneros tradicionales predominan palabras asociadas al afecto y los vínculos (como corazón, alma o familia), mientras que en el rock emergen términos vinculados al conflicto y la denuncia (como guerra, muerte o libertad), y en el género urbano se destacan expresiones relacionadas con lo cotidiano (como calle, barrio o gente). Esto sugiere que las diferencias entre emociones no dependen tanto de las palabras utilizadas, sino del contexto en el que estas adquieren significado.

4.3.3 Sentimientos

Dado que la clasificación de sentimientos contempla cincuenta categorías y presentó los menores niveles de concordancia con la referencia experta, los resultados de esta sección deben interpretarse con especial cautela. El análisis se centra en la evolución temporal de los cinco sentimientos más frecuentes identificados por el modelo: melancolía, tristeza, amor, nostalgia y resignación (ver figura 4).

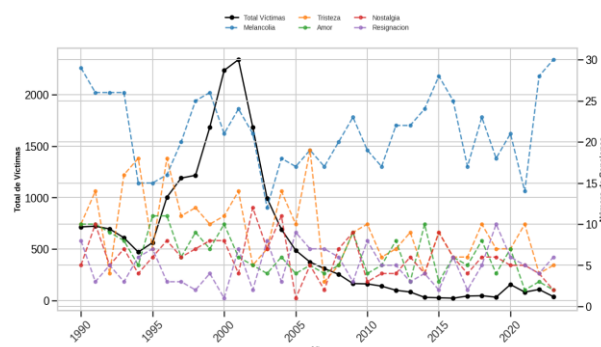


Figura 4: Sentimientos (cinco predominantes) vs total de víctimas.

La melancolía (665 letras) constituye el sentimiento predominante y mantiene una presencia relativamente estable durante el periodo analizado. La tristeza (292), el amor (193), la nostalgia (177) y la resignación (172) presentan menores frecuencias, aunque la tristeza y la nostalgia registran incrementos hacia mediados de la década de 1990 e inicios de los años 2000.

La correlación de Spearman (tabla 3) permite complementar estas observaciones. Entre los sentimientos analizados, la tristeza ($\rho \approx 0,46$; $p < 0,01$) y el amor ($\rho \approx 0,43$; $p < 0,05$) presentan asociaciones positivas significativas con el número de víctimas, mientras que la nostalgia muestra una correlación positiva moderada ($\rho \approx 0,37$; $p < 0,05$). Por el contrario, la resignación ($\rho \approx -0,15$; $p > 0,05$) y la melancolía ($\rho \approx -0,05$; $p > 0,05$) no evidencian asociaciones estadísticamente significativas.

Sentimiento	Correlación de Spearman	
	Coef	p- value
Tristeza	0,4639643	0,0057113
Amor	0,4292739	0,0112942
Nostalgia	0,3681705	0,032163
Resignación	-0,1463236	0,4089457
Melancolía	-0,0507814	0,7754815

Tabla 3: Correlaciones significativas entre sentimientos y total de víctimas por masacres.

De manera complementaria, sentimientos como la esperanza ($\rho \approx -0,60$; $p < 0,01$) y el vacío ($\rho \approx -0,36$; $p < 0,05$) presentan asociaciones negativas significativas. El detalle completo de las correlaciones para las cincuenta categorías analizadas se presenta en la tabla 6 anexo 3.

El análisis por género muestra una mayor presencia de sentimientos asociados al dolor, como melancolía, tristeza y resignación, en la mayoría de los géneros musicales. En el rock y

la música urbana aparecen con mayor frecuencia categorías relacionadas con la crítica y el conflicto, mientras que, en el folclor, la música popular y regional latinoamericana y la música comercial se observa una presencia relativamente mayor de sentimientos vinculados al afecto, como el amor y la esperanza.

A nivel léxico, se identifican términos recurrentes como “vida” y “amor” en la mayoría de los géneros analizados. Sin embargo, los patrones observados sugieren que las diferencias entre categorías emocionales podrían estar asociadas no solo al vocabulario empleado, sino también a los contextos en los que dichas expresiones adquieren significado.

5 Conclusiones

Este estudio exploró la relación entre el contenido emocional de letras de canciones producidas en Colombia entre 1990 y 2023 y la coyuntura sociopolítica del periodo, aproximada mediante el número anual de víctimas de masacres. Para ello, se aplicaron técnicas de Procesamiento de Lenguaje Natural sobre un *corpus* de 3.400 canciones, clasificadas en términos de polaridad, emociones y sentimientos mediante Modelos de Lenguaje de Gran Tamaño (LLM).

Los resultados muestran un predominio de la polaridad negativa, con una fuerte presencia de emociones y sentimientos asociados al dolor, particularmente la tristeza y la melancolía. No obstante, los coeficientes de correlación fueron, en su mayoría, bajos y no significativos, lo que sugiere que las dinámicas emocionales observadas en las letras no pueden explicarse exclusivamente a partir de los ciclos de violencia asociados al conflicto armado.

No obstante, algunas categorías específicas presentaron asociaciones estadísticamente significativas. La tristeza, la nostalgia y el amor mostraron correlaciones positivas de magnitud moderada, mientras que la ira y la esperanza presentaron asociaciones negativas significativas. Estos resultados sugieren la existencia de patrones emocionales diferenciados dentro del corpus analizado, aunque su interpretación debe realizarse con cautela debido a las limitaciones inherentes al proceso de clasificación automática.

El análisis por géneros evidenció diferencias en la distribución de emociones y

sentimientos. Mientras el folclor, la música popular y regional latinoamericana y la música comercial mostraron una mayor presencia de categorías asociadas al afecto y las relaciones personales, el rock y la música urbana presentaron una mayor frecuencia de categorías vinculadas a la crítica, el conflicto y la inconformidad social. En conjunto, estos hallazgos sugieren que la música constituye un espacio de expresión simbólica en el que convergen experiencias individuales, vínculos afectivos y elementos asociados a contextos sociales más amplios.

Finalmente, este estudio debe entenderse como una aproximación exploratoria al análisis computacional de contenidos emocionales en la música colombiana. Los resultados obtenidos muestran el potencial de los Modelos de Lenguaje de Gran Tamaño para abordar este tipo de fenómenos culturales, pero también ponen de manifiesto limitaciones importantes, particularmente en la clasificación de sentimientos específicos, donde se observaron menores niveles de concordancia con la referencia experta.

Futuras investigaciones podrían profundizar en géneros musicales específicos, incorporar otras variables relacionadas con el conflicto armado –como desplazamiento forzado, secuestro o desaparición– y continuar explorando estrategias metodológicas orientadas a mejorar la clasificación automática de contenidos emocionales. Entre estas alternativas se encuentran la clasificación jerárquica, la agrupación de categorías semánticamente cercanas (en el caso de los sentimientos) y las técnicas de ajuste especializado (fine-tuning), aplicables en contextos donde se disponga de conjuntos de datos etiquetados más amplios. Asimismo, la participación de múltiples expertos en el proceso de anotación permitiría fortalecer la robustez de futuras evaluaciones y mejorar la validación de los resultados obtenidos.

Bibliografía

Albán Morales, M. F., & Gualoto Fuentes, B. X. (2024). *Desarrollo de un algoritmo para el análisis de sentimientos de textos en kichwa en el ámbito ecuatoriano* [Trabajo de titulación, Universidad Politécnica Salesiana]. Repositorio institucional Universidad Politécnica Salesiana.

- <https://dspace.ups.edu.ec/bitstream/123456789/27234/1/TTS1768.pdf>
- Álvarez, J. (2024, 11 de noviembre). Diferencia entre emociones y sentimientos. *Gabinete de Psicología Málaga*. <https://gabinetede-psicologiamalaga.com/diferencia-emociones-sentimientos/>
- Becerra González, F. A. (2022). *El mundo y Colombia ante mis ojos: Imaginarios y resistencias socioculturales del metal en las trayectorias de Masacre y La Pestilencia (1970-2005)* [Tesis de maestría]. Repositorio Universidad Nacional. <https://repositorio.unal.edu.co/handle/unal/84477>
- Castro Rueda, J. F., & Quishpe Gaibor, J. S. (2019). Influencia de la música actual en la ética de la sociedad. *Revista Caribeña de Ciencias Sociales (RCCS)*, 11. <https://dialnet.unirioja.es/servlet/articulo?codigo=9075590>
- Dirección Nacional de Derecho de Autor. (2024). *Sobre DNDA | Dirección Nacional de Derechos de Autor*. DNDA. <https://www.derechodeautor.gov.co/es/sobre-dnda/quienes-somos/sobre-dnda>
- Flores, É. C., & Soriano, J. A. H. (2019). Análisis de concordancia mediante coeficiente de Kappa de Cohen para la elaboración de un instrumento de categorización de entrevistas biográfico-narrativas. *Educación: aportaciones metodológicas*, 50-70.
- García Vázquez, O., Alcántara, T., Sidorov, G., & Calvo, H. (2023). Clasificación automática de sentimientos en textos de canciones en idioma español. *Research in Computing Science*, 152(8), 229-237. https://www.rcs.cic.ipn.mx/2023_152_8/Clasificacion%20automatica%20de%20sentimientos%20en%20textos%20de%20canciones%20en%20idioma%20espanol.pdf
- Han, X., Chen, F., & Ban, J. (2023). Music Emotion Recognition Based on a Neural Network with an Inception-GRU Residual Structure. *Electronics*, 12(4), 978. <https://doi.org/10.3390/electronics12040978>
- Hormigos, J. (2010). La creación de identidades culturales a través del sonido. *Comunicar: Revista Científica de Comunicación y Educación*, 17(34), 91-98. <https://dialnet.unirioja.es/servlet/articulo?codigo=3164958>
- Leperski, K. (2017). El paradigma de las emociones básicas y su investigación. Hacia la construcción de una crítica [Ponencia]. *IX Congreso Internacional de Investigación y Práctica Profesional en Psicología; XXIV Jornadas de Investigación; XIII Encuentro de Investigadores en Psicología del MERCOSUR*, Buenos Aires, Argentina. <https://www.aacademica.org/000-067/146.pdf>
- Muñoz López, C. A. (2021). La música como elemento de reparación integral en el postconflicto armado, caso El Salado, Colombia. *Revista Estudios Socio-Jurídicos*, 23(2), p. 61-99. <https://doi.org/10.12804/revistas.urosario.edu.co/sociojuridicos/a.9515>
- Palomeque, M., & de Lucio, J. (2021). El sentimiento de las letras de las canciones y su relación con las características musicales. *Sociedad Española para el Procesamiento del Lenguaje Natural*, 67, 95-102. <https://doi.org/10.26342/2021-67-8>
- Patil, S., Patil, R., Goudar, S., Sangani, S., & Goudar, R. H. (2025). Review on music emotion analysis using machine learning: Technologies, methods, datasets, and challenges. *Discover Applied Sciences*, 7(7), 692. <https://doi.org/10.1007/s42452-025-07178-9>
- Paul, E. (2013). *El rostro de las emociones*. RBA Libros
- Rabino, D. N. (2024). Análisis de sentimientos en la literatura de tema medieval contemporánea. *Scriptorium*, 10(38), 121-149. <http://scriptorium.com.ar/wp-content/uploads/2024/05/SCRIPTORIUM-n38-2024.pdf>
- Sobrino Sande, J. C. (2018). *Análisis de sentimientos en Twitter*. [Tesis de maestría]. Repositorio Universitat Oberta de Catalunya. <http://hdl.handle.net/10609/81435>
- Srinilta, C., Sunhem, W., Tungjitnob, S., Thasanthiah, S., & Vatathanavaro, S. (2017). Lyric-based Sentiment Polarity Classification of Thai Songs. *Proceedings of the International MultiConference of Engineers and Computer Scientists*. 1, 5. https://www.iaeng.org/publication/IMECS2017/IMECS2017_pp364-368.pdf
- Susatama Ruiz, A. M., Hernández Divantoque, C. A., & Calderón Pérez, C. F. (2022). ¿Amor u odio en Twitter? Análisis de

Sentimientos del Banco de Bogotá basado en Procesamiento de Lenguaje Natural. [Tesis de maestría]. CRAI Universidad Central.

https://ucentral.primo.exlibrisgroup.com/permalink/57UCEN_INST/1jivd1m/alma99862075404736

Tagg, P. (2012). *Music's Meanings: A modern musicology for non-musos.* The Mass Media Music Scholars' Press.

Velázquez Carrillo, M. (2021). *Identificación de la polaridad de un texto dada la emoción de su autor* [Reporte del Proyecto de Integración]. Universidad Autónoma Metropolitana.

<https://zaloamati.azc.uam.mx/server/api/core/bitstreams/ab28ea08-97b1-4991-a90a-28b3bc40331a/content>

A Anexo 1: Tabla de resultados

Polaridad					
	Asertividad	Precisión	Recall	F1-Score	Cohen's Kappa
Gemma 4B	43,8	0.6986	0.4985	0.5803	0.2755
Gemma 12B	60,5	0.7477	0.5410	0.6152	0.3255
Gemma 27B	43	0.5827	0.5228	0.4958	0.1515
DeepSeek R1 8B	48,3	0.7193	0.6809	0.6985	0.4686
DeepSeek R1 14B	66,9	0.7338	0.5532	0.6294	0.3421
DeepSeek R1 32B	66,9	0.7352	0.7295	0.7323	0.5325
Emoción					
Gemma 4B	45,5	0.6228	0.4573	0.5079	0.2390
Gemma 12B	43,9	0.6228	0.4451	0.4947	0.2390
Gemma 27B	50,9	0.6341	0.5122	0.5439	0.2953
DeepSeek R1 8B	52,4	0.6131	0.5305	0.5635	0.3130
DeepSeek R1 14B	56,1	0.6207	0.5701	0.5882	0.3475
DeepSeek R1 32B	56,7	0.6256	0.5762	0.5930	0.3651
Sentimientos					
Gemma 4B	3,3	0.2102	0.0536	0.0655	0.0396
Gemma 12B	4,2	0.4137	0.1012	0.1135	0.0765
Gemma 27B	7	0.2682	0.1369	0.1497	0.0985
DeepSeek R1 8B	5,2	0.1347	0.0982	0.0954	0.0696
DeepSeek R1 14B	10,3	0.2633	0.1518	0.1776	0.1176
DeepSeek R1 32B	6,1	0.2678	0.1399	0.1503	0.1131

Tabla 4: Resultados de los modelos LLM en la clasificación de polaridad, emociones y sentimientos.

B Anexo 2: Letras por género musical

Género	Cantidad de letras
Folclor y tradición colombiana	1079
Urbano	634
Rock	597
Música popular y regional latinoamericana	556
Música comercial, temática y moderna	534

Tabla 5: Cantidad de letras por género.

C Anexo 3: Listado completo de correlaciones de sentimientos y total de víctimas de masacres.

Sentimiento	Correlación de Spearman	
	Coef	p- value
Esperanza	-0,6006162	0,0001729
Tristeza	0,4639643	0,0057113
Amor	0,4292739	0,0112942
Nostalgia	0,3681705	0,032163
Vacío	-0,3656656	0,0334452
Ira	-0,3129986	0,0714814
Pesar	0,3002437	0,0844768
Éxtasis	-0,2803274	0,1083134
Alegría	0,2680819	0,1252962
Compasión	0,2601581	0,1372919
Coraje	-0,253799	0,1475126
Conmoción	0,2484299	0,1565642
Perplejidad	0,2349793	0,1809868
Aceptación	-0,2050844	0,2446159
Susto	-0,2040674	0,2470144
Agradecimiento	-0,2040674	0,2470144
Frustración	-0,1978551	0,2620055
Encanto	0,1667919	0,3457807
Entusiasmo	0,1642098	0,3534047
Dolor	0,156081	0,3780587
Gozo	0,1511741	0,3934166
Desinterés	0,1508324	0,3944991
Miedo	0,1483761	0,4023326
Resignación	-0,1463236	0,4089457
Indiferencia	-0,145634	0,4111812
Desesperanza	-0,1404975	0,4280492
Sorpresa	0,1231177	0,4878785
Resistencia	-0,1153425	0,5159679
Desagrado	0,1023611	0,5645788
Rabia	-0,0946963	0,5942302
Felicidad	0,0928508	0,6014699
Intriga	-0,08375	0,6377087
Sufrimiento	0,0778752	0,6615546
Paz	-0,0778309	0,6617357
Sentimentalismo	-0,0689531	0,6983971
Deseo	-0,0551572	0,7566942
Desdicha	-0,0540473	0,761447
Orgullo	0,0527899	0,7668416
Melancolía	-0,0507814	0,7754815
Estremecer	-0,0475668	0,7893642
Reflexión	-0,0443628	0,8032647
Culpa	0,034808	0,8450555
Curiosidad	0,0266175	0,8812173
Afecto	0,0266175	0,8812173
Furia	-0,0252488	0,8872851
Odio	0,0132647	0,9406477
Shock	0,0129564	0,9420246
Enamoramiento	-0,007465	0,9665781
Resentimiento	-0,0053177	0,9761886

Depresión	-0,0023873	0,989309
-----------	------------	----------

Tabla 6: Correlaciones de sentimientos y total de víctimas de masacres.