



Caracterización espectral de música colombiana: Vallenato y Currulao.

Katherine Guerrero Caicedo
Suranyi Hurtado Campaz

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2025

Caracterización espectral de música colombiana: Vallenato y Currulao.

Katherine Guerrero Caicedo
Suranyi Hurtado Campaz

Trabajo de grado presentado como requisito parcial para optar al título de:
Estadística

Director:
Ph.D. Cesar Andres Ojeda Echeverry
Codirectora:
Ph.D. Luz Adriana Pereira Hoyos

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2025

Agradecimientos

“Entrad por sus puertas con acción de gracias, por sus atrios con alabanza; alabadle, bendecid su nombre. Porque el Señor es bueno; para siempre es su misericordia, y su fidelidad por todas las generaciones.” Salmo 100:4-5

Agradecemos a Dios por ser nuestra guía y fortaleza en este camino. A nuestras familias, por su amor incondicional y apoyo constante, y a cada una de nosotras por el compromiso mutuo y el esfuerzo compartido para alcanzar esta meta.

Resumen

Este trabajo estudia la densidad espectral en señales musicales de los ritmos tradicionales Vallenato y Currulao de Colombia, empleando herramientas avanzadas de procesamiento de señales con el fin de identificar posibles patrones y estructuras en la densidad espectral. Se aplicaron técnicas como el método de Welch, utilizando periodogramas para identificar características espectrales clave y frecuencias dominantes en las composiciones de ambos ritmos. Los resultados espectrales permitieron clasificar las señales musicales y comprender las diferencias estructurales entre los dos géneros. Además, se implementaron máquinas de soporte vectorial (SVM) para realizar la clasificación de las canciones. Los resultados resaltan la utilidad del análisis espectral en la caracterización de géneros musicales.

Palabras clave: Densidad espectral, procesamiento de señales, método de Welch, periodogramas, señales musicales, características espectrales, frecuencias dominantes, clasificación musical, géneros musicales, algoritmos de clasificación, máquinas de soporte vectorial, inteligencia artificial.

Abstract

This work studies the spectral density in musical signals of the traditional Vallenato and Currulao rhythms of Colombia, employing advanced signal processing tools in order to identify possible patterns and structures in spectral density. Techniques such the Welch method were applied, using periodograms to identify key spectral features and dominant frequencies in the compositions of both rhythms. The spectral results allowed us to classify the musical signals and understand the structural differences between the two genres. In addition, vector support machines (SVM) were implemented to classify the songs. The results highlight the usefulness of spectral analysis in the characterization of musical genres.

Contenido

Lista de Figuras	VII
Lista de Tablas	VIII
1. Introducción	1
1.1. Planteamiento del Problema	2
1.2. Justificación.	4
1.3. Pregunta de investigación	4
2. Objetivos	5
2.1. Objetivo General	5
2.2. Objetivos Específicos	5
3. Antecedentes.	6
3.1. Antecedentes contextuales	6
3.2. Antecedentes Estadísticos	7
4. Marco Teórico	9
4.1. Géneros musicales	9
4.2. Series de tiempo en el análisis espectral	9
4.2.1. Teoría de procesamiento de señales y análisis espectral	10
4.2.2. Conceptos relevantes en las series de tiempo	11
4.3. Análisis de espectro usando el método de Welch	13
4.3.1. Método de Welch	13
4.4. Características espectrales	14
4.5. Clasificación de géneros musicales basada en característica espectrales	16
4.5.1. Clasificación con Random Forest (RF)	16
4.5.2. Clasificación con Redes Neuronales (RNA)	17
4.5.3. Clasificación con Máquinas de Soporte Vectorial (SVM)	17
5. Metodología	19
5.1. Recolección de Datos	19
5.2. Instalación y Configuración del Entorno	19
5.3. Configuración de Directorios	19
5.4. Extracción de Características del Audio	20
5.5. Estimación de la densidad espectral	20
5.5.1. Procesamiento de Señales	20
5.5.2. Aplicación del Método de Welch	20

5.6.	Análisis para identificar patrones y estructuras	21
5.7.	Selección de características	21
5.8.	Clasificación binaria de generos musicales	22
5.8.1.	Clasificación mediante SVM	23
5.8.2.	Clasificación de Nuevas Canciones	23
6.	Resultados	25
6.1.	Descripción de los Datos	25
6.2.	Análisis de Series Temporales	25
6.3.	Análisis de densidades	28
6.4.	Características generales y espectrales	31
6.5.	Clasificación	35
6.5.1.	Comparación de tres algoritmos	35
6.5.2.	Mejor algoritmo	35
6.5.3.	Selección de características	36
6.5.4.	Experimentación con maquinas de soporte vectorial	36
6.5.5.	Clasificación por medio de maquina de soporte vectorial	39
6.5.6.	Implementación de clasificación a nuevas canciones	41
7.	Conclusiones y Recomendaciones	42
7.1.	Conclusiones	42
7.2.	Recomendaciones	43
A.	Anexo	47

Lista de Figuras

6-1. Espectrograma del Currulao.	26
6-2. Espectrograma del Vallenato.	26
6-3. Comparación del periodograma promedio por género.	27
6-4. Distribución de potencia en frecuencia 10.	29
6-5. Densidad de energía por género.	30
6-6. Importancia de las características en las maquinas de soporte vectoriales.	37

Lista de Tablas

6-1.	Características generales de las señales del género Currulao.	31
6-2.	Frecuencias específicas de las señales del género Currulao.	32
6-3.	Características generales de las señales del género Vallenato.	33
6-4.	Frecuencias específicas de las señales del género Vallenato.	34
6-5.	Comparación de Modelos de Clasificación para ambos géneros musicales	35
6-6.	Diferencias clave entre las características de Currulao y Vallenato	36
6-7.	Comparación de precisión entre modelos SVM	38
6-8.	Resultados de la prueba de igualdad de proporciones	38
6-9.	Matriz de confusión	39
6-10.	Métricas de Evaluación del Modelo	40
6-11.	Clasificación de canciones	41

1. Introducción

La música es una expresión cultural y artística, que ha sido objeto de estudio desde múltiples perspectivas. En la actualidad, los avances en el procesamiento digital de señales y el aprendizaje automático han abierto nuevas alternativas, transformado significativamente la manera de estudiar los datos multimedia, particularmente las características de las señales de audio. Las señales de audio pueden representarse a través de señales dependientes del tiempo, por esto se podrían catalogar como series temporales que contienen información sobre fenómenos acústicos que pueden ser analizados mediante metodologías estadísticas. (Shumway and Stoffer, 2017).

El análisis de series de tiempo para señales de audio logra descomponer el contenido de las señales en componentes fundamentales, permitiendo la identificación de patrones y tendencias. En este contexto, técnicas como la transformada de Fourier y el análisis espectral cobran alta relevancia, ya que logran representar señales en el ámbito de la frecuencia, identificando las componentes regulares que caracterizan a una muestra de audio (Martín et al., 2001). Estas metodologías no solo son útiles para el estudio de fenómenos físicos, sino que también han demostrado efectividad en aplicaciones como la clasificación automática de audio, el reconocimiento de voz, y la caracterización de géneros musicales (Tzanetakis and Cook, 2002).

El análisis espectral a través del método de Welch ha sido ampliamente utilizado debido a su capacidad para reducir la varianza en la estimación de la densidad espectral de potencia, mediante el promedio de periodogramas segmentados y solapados (Welch, 1967). Esta propiedad lo convierte en una herramienta robusta para el análisis de señales complejas, como las señales de audio, donde la presencia de múltiples frecuencias simultáneas requiere métodos estadísticos confiables.

Al mismo tiempo, la integración de técnicas estadísticas con modelos de aprendizaje automático ha mejorado las posibilidades para clasificar automáticamente muestras musicales según sus características espectrales. Entre estas técnicas, las Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) han demostrado ser eficientes en trabajos de clasificación supervisada. (Cortes and Vapnik, 1995).

Aunque hay un crecimiento en la investigación en clasificación musical, son pocos los estudios enfocados en géneros tradicionales colombianos como el Vallenato y el Currulao, estos poseen características estructurales distintivas que podrían ser capturadas mediante descriptores espectrales, lo que significa una oportunidad de análisis desde una perspectiva estadística.

En este trabajo se propone una metodología exploratoria orientada a la caracterización y clasificación de los géneros musicales colombianos Vallenato y el Currulao a partir de sus

1 Introducción

representaciones espectrales. Para ello, se emplea el método de Welch como herramienta de análisis espectral de señales musicales, y se desarrolla un modelo de clasificación automática mediante SVM, utilizando como insumo principal las características derivadas del espectro de potencia. El estudio tiene como objetivo principal evaluar la efectividad de estas representaciones en la discriminación automática de estos dos géneros musicales.

La propuesta metodológica se desarrolla empleando el software estadístico R, para el análisis espectral como para la implementación de modelos de aprendizaje automático. Este enfoque no solo permite una aproximación cuantitativa robusta, sino que también contribuye a la construcción de herramientas de análisis reproducibles, orientadas al estudio de expresiones musicales tradicionales desde una perspectiva estadística.

1.1. Planteamiento del Problema

Colombia es un país caracterizado por su gran diversidad y riqueza cultural, fruto de la mezcla de culturas africanas, europeas e indígenas. Esta diversidad se refleja en manifestaciones culturales como los géneros musicales tradicionales, los cuales varían según la región. Así, en la región andina predominan ritmos como el bambuco y el pasillo; en la región caribe, la cumbia y el vallenato; en la región pacífica, el currulao; y en la Orinoquía, el joropo, entre otros (Escobar, 2005).

En particular, el Currulao, típico de la región pacífica, y el Vallenato, originario de la región caribe, son géneros musicales que han mantenido su vigencia a lo largo del tiempo y siguen siendo importantes símbolos de identidad cultural en sus respectivas regiones (Londoño, 1985; Aponte Mantilla, 2011).

Sin embargo, la globalización y la influencia de fenómenos culturales contemporáneos han provocado que muchos de estos géneros tradicionales pierdan presencia, reconocimiento y difusión, especialmente en plataformas digitales (Cruz et al., 2019).

A pesar de que en el mundo científico haya una creciente ola por investigar las señales de audio, más específicamente la caracterización y clasificación de géneros musicales como el rock, reggae, salsa o música clásica a través del estudio de sus características espectrales para dicho fin, en el caso de géneros musicales tradicionales como el vallenato y el currulao este tipo de investigaciones es casi nula (Academia Lab, 2025).

Ante todo lo anterior surge la necesidad de aplicar metodologías estadísticas para el análisis de las señales de audio para estos géneros musicales con el propósito de identificar características que permitan caracterizarlos, con el fin de poder clasificarlos entre sí. En ese sentido, el método de Welch, que es una técnica para estimar la densidad espectral de potencia en diferentes frecuencias, se basa en utilizar estimaciones del espectro de periodograma que resultan de convertir una señal del dominio del tiempo al dominio de la frecuencia (Academia Lab, 2025).

En términos generales, en este trabajo el problema que se plantea es la falta de caracterizaciones

estadística para géneros musicales tradicionales colombianos partiendo del análisis espectral de señales de audio. Por medio del método de Welch se quiere analizar las frecuencias del vallenato y el currulao para así identificar características, patrones repetitivos que permitan tener diferencias entre ambos géneros para así clasificarlos, contribuyendo de esta manera en un estudio de los mismos desde un punto de vista numérico y estadístico.

1.2. Justificación.

La caracterización espectral de los géneros musicales tradicionales: vallenato y currulao es un reto tanto en lo académico como en lo tecnológico. Aunque ambos han sido objeto de estudio de manera cultural e histórica, en el ámbito estadístico hay una exploración limitada de estos géneros musicales que contribuya a su caracterización y clasificación a partir de su densidad espectral.

Es así como este trabajo se justifica en la necesidad de implementar un estudio basado en el procesamiento de las señales de audio de ambos géneros musicales y en métodos de aprendizaje automático para el análisis de las características acústicas más representativas de estos géneros musicales. Por lo cual, se utiliza el método de Welch para la estimación de la densidad espectral de potencia, lo que permite tener de manera detallada la energía distribuida en las diferentes frecuencias en la señal de audio, lo que facilita la identificación de características que pueden ser utilizadas en la clasificación.

También como parte de la metodología se hará uso de las máquinas de soporte vectorial para la clasificación de los fragmentos musicales en cada una de las categorías que en este caso son nuestros dos géneros musicales. La elección de este método de clasificación se debe a lo efectivo que resulta en la clasificación con datos de alta dimensión y en su efectividad para encontrar divisiones entre las clases.

En términos generales, el enfoque que se ha propuesto en este trabajo de grado permite analizar de manera estadística los dos géneros musicales y contribuir a un modelo de clasificación basado en las características de estos. Además, es un modelo que puede ser replicado y extendido a otros géneros musicales. Lo que contribuye al desarrollo de metodologías en investigaciones académicas enfocadas en la preservación de patrimonio musical.

1.3. Pregunta de investigación

En el presente trabajo de investigación se plantea la siguiente pregunta.

¿Cómo utilizar métodos de caracterización espectral y algoritmos de clasificación, como el método de Welch y las máquinas de soporte vectorial, para identificar patrones distintivos en las señales de audio del vallenato y el currulao para llegar a clasificarlos?

2. Objetivos

2.1. Objetivo General

Estudiar la densidad espectral de los géneros musicales tradicionales Vallenato y Currulao de Colombia.

2.2. Objetivos Específicos

- Estimar la densidad espectral de las canciones de los ritmos musicales Vallenato y el Currulao.
- Realizar un análisis a fin de identificar posibles patrones y estructuras en la densidad espectral de los ritmos musicales Vallenato y Currulao.
- Proponer una regla de clasificación para los ritmos musicales estudiados a partir de las densidades espectrales.

3. Antecedentes.

3.1. Antecedentes contextuales

Debido a fenómenos como la globalización, los avances tecnológicos y el crecimiento de la industria musical, la investigación enfocada en el análisis de géneros musicales ha aumentado significativamente en los últimos años. Esto se evidencia en trabajos como el desarrollado en la Universidad Nacional de Colombia (sede Bogotá), donde en 2019 se publicó el artículo “Automatic Identification of Traditional Colombian Music Genres Based on Audio Content Analysis and Machine Learning Techniques” Cruz et al. (2019).

En este estudio, se propuso un modelo de aprendizaje automático para clasificar seis géneros musicales tradicionales colombianos (bambuco, carranga, cumbia, joropo, pasillo y vallenato), utilizando un conjunto de datos de 180 piezas musicales. Se emplearon metodologías de clasificación supervisada (Random Forest y redes neuronales) y no supervisada (clustering), así como técnicas de reducción de dimensionalidad (PCA y t-SNE). El modelo fue implementado en Google Colaboratory, utilizando librerías de Python como Scikit-learn y Keras.

Entre los resultados obtenidos, se destaca un 69 % de precisión general en la clasificación de géneros. Además, se observó que el método de Random Forest tuvo un rendimiento superior en un 8 % respecto al método de clustering. Este trabajo evidenció el potencial del análisis automatizado de señales de audio para el reconocimiento de géneros musicales tradicionales colombianos.

Por otra parte, en el artículo “Music Genre Classification Using Spectral Analysis and Sparse Representation of the Signals” desarrollado en la University of British Columbia (Canadá) Banitalebi-Dehkordi and Banitalebi-Dehkordi (2014), se abordó la clasificación de diez géneros musicales internacionales (clásico, country, disco, hip-hop, jazz, rock, blues, reggae, pop y metal) mediante un método novedoso de extracción de características. Este método combinó información a corto plazo (análisis de ventanas de señal) y a largo plazo (combinación de características temporales), logrando una mayor eficiencia computacional en comparación con otros enfoques tradicionales como el muestreo por compresión. Los aportes de este estudio residen en optimizar la representación de señales musicales para su clasificación automática mediante técnicas de análisis espectral.

Un enfoque diferente se encuentra en el análisis de señales biológicas, como se expone en el libro Sound Analysis and Synthesis with R de Jérôme Sueur Sueur et al. (2018). Aunque el objetivo de este trabajo fue la caracterización acústica de insectos y aves, su metodología basada en el análisis espectral de señales sonoras es altamente aplicable al estudio de sonidos musicales. La capacidad

de caracterizar entidades biológicas a partir de señales acústicas, sin intervención directa en su entorno, demuestra el potencial del análisis espectral como herramienta científica en el estudio de datos no estructurados como el sonido.

En otro ámbito, en Echeverry et al. (2007) desarrollaron un análisis armónico de señales biomédicas (microelectrodos de registro - MER) aplicando el método de Welch para calcular la densidad espectral de potencia. El objetivo fue caracterizar las componentes de frecuencia de distintas zonas cerebrales en pacientes con enfermedad de Parkinson. Para optimizar el análisis, aplicaron técnicas de ventaneo y clasificadores lineales. Su metodología permitió identificar eficazmente 66 señales correspondientes a distintas zonas cerebrales (41 de la sustancia negra y 25 del núcleo subtalámico), destacando la relevancia del análisis espectral en la caracterización de señales no estructuradas en el ámbito médico.

Estos antecedentes reflejan el amplio campo de aplicación de las metodologías de análisis de señales, en especial el análisis espectral, tanto en el área musical como en otras disciplinas, reafirmando la pertinencia de su aplicación en la caracterización de géneros musicales tradicionales como el Vallenato y el Currulao.

3.2. Antecedentes Estadísticos

El Análisis Espectral Singular (SSA, por sus siglas en inglés) es una técnica estadística que se utiliza principalmente en el estudio de series temporales no estacionarias. Su principal utilidad está radicada en la descomposición de las señales en sus componentes interpretables como tendencias, ruidos y oscilaciones periódicas, lo que permite realizar un análisis flexible y más detallado a los datos. Esta metodología ha venido siendo aplicada con éxito en diversos campos, debido a su buena capacidad para capturar estructuras presentes en las series de tiempo. (Carvalho and Silva, 2012)

En Carvalho and Silva (2012) se aplicó el SSA para estudiar las fluctuaciones del ciclo económico en Estados Unidos, centrándose en el análisis de la brecha de producción. Esta investigación propuso realizar una descomposición de las series temporales en componentes clave que ayudan a identificar con mayor claridad los patrones subyacentes del comportamiento económico. Los autores encontraron que, en comparación con métodos tradicionales como los filtros, el SSA generaba una mayor precisión y flexibilidad en la descomposición de las señales.

Un resultado relevante de ese estudio fue la capacidad del SSA para identificar variaciones económicas de largo plazo con un alto nivel de detalle. Por ejemplo, al utilizar una ventana de análisis de 32 trimestres, logró captar patrones amplios sin que se fugaran aspectos significativos de corto plazo. Este enfoque demostró ser bastante útil para entender más las dinámicas del ciclo económico en situaciones reales.

El proceso del SSA tiene dos fases principales: la descomposición y la reconstrucción. En la primera fase, la serie temporal inicial se transforma en una matriz, la cual es dividida en componentes principales que reflejan las fluctuaciones más sobresalientes de la serie. En la segunda fase, la de

3 Antecedentes.

reconstrucción, se eligen los componentes significativos y se combinan para crear una nueva versión filtrada de la serie, donde se resalta sus estructuras de mayor importancia. (Carvalho and Silva, 2012).

Uno de los aspectos fundamentales en la aplicación del SSA es la elección de la longitud de la ventana para la fase de descomposición, ya que esta elección tiene un impacto directo en los componentes extraídos y, por lo tanto, impacta en la precisión del análisis. (Carvalho and Silva, 2012).

4. Marco Teórico

En este apartado se definen los conceptos esenciales para el desarrollo de este trabajo. Se inicia por la definición de los géneros musicales; seguidamente se enuncian conceptos relacionados con series de tiempo, específicamente con el análisis espectral, el método Welch, características espectrales e implementación de un modelo de clasificación de géneros musicales basado en características espectrales.

4.1. Géneros musicales

El género musical es una herramienta categórica que clasifica y organiza las piezas musicales basándose en similitudes entre características como la estructura rítmica, la instrumentación y la composición de las obras. Los géneros son esenciales para definir y organizar las colecciones musicales (Scaringella et al., 2006).

Vallenato

El Vallenato es un género musical tradicional surgido de la fusión de expresiones culturales del norte de Colombia: canciones de los vaqueros del Magdalena Grande, cantos de los esclavos africanos y ritmos de danzas tradicionales de los pueblos indígenas de la Sierra Nevada de Santa Marta (UNESCO, 2015).

Currulao

El Currulao es un ritmo musical colombiano, desarrollado en la región pacífica. Este género musical se originó por personas esclavizadas a finales de la época colonial. Su nombre proviene del tambor Currulao y del tambor conuno, con los cuales se acompaña la melodía del mismo (Escobar, 2005).

4.2. Series de tiempo en el análisis espectral

El análisis de señales constituye un campo fundamental dentro de la ingeniería y las ciencias aplicadas, ya que permite estudiar fenómenos complejos a través de su representación matemática en el dominio del tiempo, la frecuencia o ambos. En particular, el estudio de señales sonoras, como las que conforman piezas musicales, requiere herramientas que permitan identificar patrones relevantes y características estructurales que no siempre son evidentes en el dominio temporal.

La música, entendida como una señal acústica, puede ser abordada como una serie temporal cuyos componentes varían con el tiempo de forma compleja y no siempre predecible. Por esta razón, las

metodologías de procesamiento de señales y análisis espectral se convierten en instrumentos clave para su caracterización, especialmente cuando se busca identificar rasgos distintivos de diferentes géneros musicales o estilos regionales, como ocurre en este trabajo con el vallenato y el currulao.

En esta sección se presentan los fundamentos teóricos del procesamiento de señales y del análisis espectral, necesarios para comprender el enfoque metodológico adoptado en este estudio. Se introducen conceptos esenciales como la definición de señal, sistema y procesamiento, que permitirán entender cómo se realiza la extracción de información relevante a partir de datos acústicos.

4.2.1. Teoría de procesamiento de señales y análisis espectral

Una señal es una manifestación de un fenómeno expresada como una cantidad física. Desde este punto de vista, se considera que la señal aporta algún tipo de información sobre el estado, comportamiento y/o características del estímulo y del sistema físico que produjo la señal.

Una parte importante de las señales con que se trabaja en ingeniería, se genera por medios naturales como la voz, la velocidad del viento, o la humedad relativa.

Los valores que toma una señal (variable dependiente) son función de una o más variables independientes como el tiempo, la intensidad de luz, la velocidad, la temperatura y el peso. Una señal de voz es una función de la variable independiente tiempo, la intensidad de luz en cualquier punto de una imagen es una función de dos variables espaciales, mientras que un punto en una escena de vídeo es función de tres variables independientes: dos espaciales y una temporal. Los valores de la variable dependiente pueden ser reales, complejos o vectores (Mata Hernández et al., 2017).

Un sistema es una relación funcional que asigna una señal de salida $y(t)$, a una señal de entrada $x(t)$, y que puede expresarse matemáticamente como:

$$y(t_0) = [x(t); -\infty < t < \infty], \quad -\infty < t_0 < \infty \quad (4-1)$$

Un sistema puede definirse como un dispositivo físico que realiza una operación sobre una señal. La forma en la que se generan las señales se encuentra asociada con la respuesta de un sistema ante un estímulo o fuerza. El estímulo en combinación con el sistema constituye una Fuente de Señal.

Procesamiento de señal

Las señales contienen información, y normalmente cuando se va hacer uso de ellas es necesario modificarlas o transformarlas para extraer, cambiar o adaptar esa información de acuerdo a las necesidades que se necesitan en el desarrollo que se va a realizar. Por ello el procesamiento de señales se puede definir como el conjunto de operaciones preestablecidas que se realizan en una señal con el propósito de obtener información que pueda ser útil o adaptar dicha información en un

contexto específico (Ambardar, 1995).

Dependiendo de la naturaleza de la señal, el proceso de la misma se puede hacer en dominios analógicos, digitales o híbridos. En el caso cuando las señales son analógicas que van a ser procesadas digitalmente, es necesario primero transformarlas por medio de un convertidor analógico a digital (ADC). Luego de dicho procesamiento, la señal resultante puede necesitar una conversión para volver al formato analógico, en este caso utilizando un convertidor de digital a analógico (DAC), para su interpretación o reproducción final.

Por otro lado, es primordial que el proceso de muestreo cumpla con el criterio de Nyquist, lo cual implica que la frecuencia de muestreo debe ser el doble de la frecuencia máxima contenida en la señal. Si no es así, se produce el fenómeno que es conocido como aliasing, que puede distorsionar la señal muestreada e impedir que se reconstruya correctamente. (Ambardar, 1995).

4.2.2. Conceptos relevantes en las series de tiempo

Un **archivo WAV (Waveform Audio File Format)** es un formato estándar utilizado para almacenar audio digital en sistemas Windows, que fue desarrollado en conjunto entre Microsoft e IBM en 1991. Esto funciona como un contenedor RIFF, que está compuesto por fragmentos o chunks como RIFF, "fmt" "data", que permiten organizar los datos y los metadatos.

Los archivos WAV resisten muchas formas de codificación de audio, aunque habitualmente es más compatible con la codificación LPCM (Linear Pulse-Code Modulation), es decir, un audio sin comprimir, esto es una configuración utilizada para aplicaciones profesionales o donde se necesita mantener la calidad del sonido.

Ya que su estructura es simple y flexible, los archivos WAV son fáciles de interpretar por una gran variedad de dispositivos y herramientas de edición de audio.

Un **proceso estocástico** Y_t es estacionario cuando las propiedades estadísticas de cualquier secuencia finita $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ ($n \geq 1$) de componentes Y_t son semejantes a las de la secuencia $Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}$ para cualquier número entero $h = \pm 1, \pm 2, \dots$

Un proceso de **ruido blanco univariante** es una secuencia (A_t) de variables aleatorias escalares idénticas e independientemente distribuidas con media 0 y varianza σ_a^2 , lo cual suele representarse como $(A_t) \sim \text{IID}(0, \sigma_a^2)$. Cuando cada A_t sigue una distribución normal, (A_t) se denomina un proceso de Ruido Blanco Normal o Gaussiano, lo cual suele representarse como $(A_t) \sim \text{NIID}(0, \sigma_a^2)$.

La **tendencia secular o tendencia** a largo plazo de una serie es por lo común el resultado de factores a largo plazo. En términos intuitivos, la tendencia de una serie de tiempo caracteriza el patrón gradual y consistente de las variaciones de la propia serie, que se consideran consecuencias de fuerzas persistentes que afectan el crecimiento o la reducción de la misma, tales como: cambios en la población, en las características demográficas de la misma, cambios en los ingresos, en la salud,

en el nivel de educación y tecnología. Las tendencias a largo plazo se ajustan a diversos esquemas. Algunas se mueven continuamente hacia arriba, otras declinan, y otras más permanecen igual en un cierto período o intervalo de tiempo.

Densidad espectral

La densidad espectral $f(\omega)$, que en algunos casos también es llamada función espectral o espectro de potencia) caracteriza cómo la varianza total de un proceso estocástico se distribuye a lo largo de las diferentes frecuencias. Este concepto es central en el análisis espectral, ya que permite estudiar la dinámica de la serie en el dominio de la frecuencia en lugar del tiempo.

Si X_t es un proceso débilmente estacionario, con media μ y función de autocovarianza $\gamma(h) = E[(X_t - \mu)(X_{t+h} - \mu)]$, la densidad espectral está definida como **la transformada de Fourier** $\gamma(h)$:

$$f(w) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-iwh}, -\pi \leq w \leq \pi. \quad (4-2)$$

Equivalentemente, la función de autocovarianza se puede recuperar como la transformada inversa de Fourier de $f(w)$:

$$\gamma(h) = \int_{-\pi}^{\pi} e^{iwh} f(w) dw. \quad (4-3)$$

Transformada rápida de Fourier La Transformada de Fourier de una función $f(t)$ es una herramienta matemática utilizada para la descomposición de una señal compleja en sus componentes de frecuencia, y es fundamental para el análisis de señales. Se define como:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (4-4)$$

Gracias a estas relaciones se establece una equivalencia matemática entre la representación temporal y la representación frecuencial del proceso. La densidad espectral, por consiguiente, realiza la descomposición a la varianza de la serie, identifica patrones cíclicos y periódicos que no logran evidenciarse en el dominio del tiempo, generaliza la noción de análisis armónico en procesos aleatorios uniendo la transformada de Fourier y la probabilidad; además, logra conectar el análisis de señales con la teoría estadística.

El **periodograma** es un estimador no paramétrico utilizado para aproximar la densidad espectral de una serie temporal. De manera más simple, permite representar cómo se distribuye la varianza o la energía de la serie en función de la frecuencia, mostrando que ciclos u oscilaciones predominan en los datos. El periodograma se deriva de la transformada de Fourier de la serie, aunque es un estimador asintóticamente insesgado, no es consistente, pues su varianza no disminuye a medida que aumenta el número de observaciones, esto provoca que las estimaciones sean muy fluctuantes (Brillinger, 1974)

$$I_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{-it\lambda} \right|^2, \quad (4-5)$$

donde X_t representa la serie temporal y N el tamaño de la muestra.

El periodograma es interpretado como una herramienta que muestra la importancia relativa de diferentes frecuencias en la señal, permitiendo identificar tendencias cíclicas o patrones recurrentes. Pero, debido a su alta variabilidad, suele combinarse con técnicas de suavizado o promediado que estabilizan la estimación y facilitan su interpretación (Bloomfield, 1976)

4.3. Análisis de espectro usando el método de Welch

4.3.1. Método de Welch

El método de Welch es una técnica no paramétrica utilizada para estimar la densidad espectral de potencia (PSD) de una señal estacionaria en el dominio de la frecuencia. Este método fue propuesto por Peter D. Welch en 1967 como un arreglo al método que estima el periodograma, esto con el propósito de reducir la varianza de la estimación sin que se comprometiera de manera significativa la resolución en frecuencia (Welch, 1967).

Este método consiste en dividir la señal original en segmentos de una longitud fija, los cuales se superponen; se debe aplicar una función de ventana a cada uno de los segmentos para minimizar los efectos de la fuga espectral, luego se debe calcular el periodograma modificado de cada uno de los segmentos a través de la Transformada de Fourier Discreta, para finalmente promediar los periodogramas anteriormente calculados y así obtener una estimación un poco más estable de la densidad espectral de potencia (III, 2011).

Este enfoque permite una estabilidad entre resolución y varianza, puesto que cuanto mayor sea la superposición y menor la longitud de los segmentos, más suave va ser la estimación de la densidad espectral de potencia, con un poco de pérdida de resolución espectral. El método de Welch es muy utilizado en diferentes campos en donde se realiza procesamiento de señales, como en el campo bio medico, acústico y de telecomunicaciones, esto debido a lo robusto en cuanto al ruido y a su facilidad computacional (Rumsey, 2003).

Cálculo del periodograma en el método Welch

Para cada segmento k , se calcula el periodograma $P_k(f)$ utilizando la Transformada de Fourier Discreta (DFT), aplicada a una porción de la señal multiplicada por una ventana:

$$P_k(f) = \frac{1}{L} \left| \sum_{n=0}^{L-1} x_k(n) w(n) e^{-i2\pi f n} \right|^2 \quad (4-6)$$

- $P_k(f)$: Periodograma estimado para el segmento k en la frecuencia f .

- L : Longitud (en muestras) de cada segmento en que se divide la señal.
- $x_k(n)$: n -ésima muestra del k -ésimo segmento de la señal.
- $w(n)$: Función ventana aplicada a cada segmento (por ejemplo, ventana de Hamming o de Hann).
- n : Índice de tiempo discreto dentro del segmento, con $0 \leq n < L$.
- f : Frecuencia (puede expresarse en Hz o como fracción de la frecuencia de muestreo).
- i : Unidad imaginaria, donde $i^2 = -1$.

Luego, los periodogramas de todos los segmentos se promedian para obtener una estimación más robusta y suave de la densidad espectral de potencia:

$$\hat{S}(f) = \frac{1}{K} \sum_{k=0}^{K-1} P_k(f) \quad (4-7)$$

- $\hat{S}(f)$: Estimación final de la densidad espectral de potencia (PSD) en la frecuencia f .
- K : Número total de segmentos en los que se divide la señal.

4.4. Características espectrales

El procesamiento de señales cuenta con un conjunto de herramientas y métodos enfocados a analizar, transformar e interpretar todo tipo de señales. Una de las aplicaciones mas frecuentes es la obtención de características espectrales, las cuales contienen propiedades cuantitativas derivadas del comportamiento de una señal en el dominio de la frecuencia. Estas características permiten comprender mejor la estructura interna de una señal, lo que facilita su clasificación, segmentación y comparación entre señales. Para la estimación de dichas características es común el uso de métodos no paramétricos como el método de Welch, el cual consiste en dividir la señal en segmentos solapados, aplicarle una ventana a cada uno y calcular el periodograma por medio de la Transformada de Fourier Discreta, para así promediar dichos segmentos y obtener una estimación robusta de la densidad espectral de potencia. A través de esta estimación espectral, es posible extraer ciertas métricas relevantes que caracterizan distintas propiedades de la señal, características como la energía, el centroide, el ancho de banda, entre otras. Conocer este tipo de características es fundamental para identificar patrones, diferenciar distintas señales y mejorar la funcionalidad de métodos de clasificación.

A continuación, se describen las características espectrales utilizadas en la caracterización y clasificación de este trabajo:

Energía

Se puede definir la energía de una señal como la suma de los cuadrados de sus valores en cada frecuencia o tiempo. La energía proporciona una medida del contenido energético global de una señal. Normalmente en clasificación, la energía ayuda a detectar variaciones en la intensidad de las señales cuando son diferentes entre sí, así permitiendo distinguir patrones que son relevantes en las señales. Según Bishop (2006), las señales con una alta energía pueden estar relacionadas con características de clases específicas debido a que su potencia se acumula en el tiempo.

Varianza

La varianza espectral estima la dispersión de los valores de la señal con respecto a la media. La varianza es una medida estadística que indica que tan dispersos están los valores de la señal respecto a su promedio. Cuando hay una varianza alta eso significa que en los datos hay mayor variabilidad en los datos, Esto resulta útil para diferenciar clases con patrones de cambio únicos. Lo anterior es importante es problemas de clasificación en donde las diferencias en la variabilidad son distintivas en diferentes categorías (Trevor Hastie, 2009).

Simetría

La simetría espectral describe la medida en la cual una señal se comporta como un espejo teniendo como punto de referencia su eje central. En clasificación, las señales simétricas muestran estructuras ordenadas y repetitivas, mientras que cuando las señales son asimétricas muestran irregularidades o anomalías. La simetría puede ser un factor determinante para identificar y clasificar diferentes tipos de señales (Bishop, 2006).

Curtosis

La curtosis espectral mide la estructura en la cual se distribuyen los datos. La curtosis muestra la notoriedad en los picos de los datos. Cuando el nivel de curtosis es elevado indica la presencia de valores extremos o aglomeraciones muy notorias en los datos. Esto es crucial para diferenciar categorías, porque se destacan características únicas o particulares que permiten separar una clase de otra, cuando se realiza una clasificación (Trevor Hastie, 2009).

Cruces Cero

Los cruces cero se refieren al número de veces en donde una señal cruza el eje cero en un determinado periodo de tiempo. Esta propiedad es fundamental para detectar variaciones en la señal, sean lentas o rápidas, estas muestran diferentes tipos de eventos o patrones. En señales acústicas y electromagnéticas los cruces cero son muy útiles (Bishop, 2006).

Centroide

El centroide se calcula como el promedio ponderado de las frecuencias que contiene la señal. Es decir, es un índice que representa la frecuencia promedio o predominante en la señal. El centroide

es fundamental para comprender la localización promedio de los valores de la señal y puede ser útil en la clasificación de las señales teniendo en cuenta las características frecuenciales (Cristianini and Shawe-Taylor, 2000).

Ancho de Banda

El ancho de banda es el resultado de la diferencia entre la frecuencia más alta y más baja en una señal. El ancho de banda describe la extensión del espectro ocupado por la señal. Técnicamente, se puede expresar como la desviación estándar de las frecuencias respecto al centroide, mostrando la complejidad de la señal.

4.5. Clasificación de géneros musicales basada en característica espectrales

La clasificación de géneros musicales requiere la extracción de características destacadas de las señales de audio, así como la aplicación de métodos de clasificación eficaces. Entre las formas más utilizadas están los coeficientes ceptrales en las frecuencias de Mel (MFCC), reconocidos ampliamente por su habilidad para capturar propiedades perceptuales del sonido. Esta técnica fue introducida por Davis y Mermelstein en 1980 como una herramienta útil para el reconocimiento de voz, mostrando un rendimiento sobresaliente frente a otros métodos de parametrización (Davis and Mermelstein, 1980).

Una vez extraídas las características espectrales, el siguiente paso es aplicar métodos de clasificación que permitan otorgar cada fragmento de audio a un género musical específico. Dentro de este campo, sobresalen los métodos de aprendizaje automático supervisado, como el algoritmo de Random Forest y las Redes Neuronales Artificiales (RNA). Estos métodos han demostrado ser eficaces en clasificaciones complejas, esto es debido a su capacidad para modelar relaciones no lineales y manejar conjuntos de datos de alta dimensionalidad. A continuación, se presenta una descripción breve de ambos métodos

4.5.1. Clasificación con Random Forest (RF)

Random Forest es un algoritmo de aprendizaje automático que se basa en el uso de múltiples árboles de decisión, el cual fue desarrollado por Leo Breiman en el año 2001. La idea principal de este método es que al combinar las predicciones de varios árboles en lugar de depender de uno solo, se puede alcanzar un modelo más robusto, preciso y menos propenso al sobreajuste (Breiman, 2001).

Cada árbol de decisión se entrena con un subconjunto aleatorio de datos, la cual es una técnica conocida como *bagging*, lo que ayuda a reducir la varianza del modelo y mejorar su capacidad de generalizar. A pesar de sus ventajas, el Random Forest puede ser costoso en términos computacionales cuando se utilizan muchos árboles de decisión, y su rendimiento también puede verse afectado

cuando se incluyen demasiadas variables irrelevantes(Louppe, 2014).

4.5.2. Clasificación con Redes Neuronales (RNA)

Las Redes Neuronales Artificiales (rna) se inspiran en cómo funciona el cerebro humano para procesar la información. Estas son capaces de identificar patrones complejos y aprender directamente a partir de los datos, esto las hace especialmente útiles para tareas como la clasificación de géneros musicales.

Una red neuronal está compuesta por varias capas de neuronas artificiales organizadas de la siguiente manera:

- **Capa de entrada:** Es la encargada de recibir los primeros datos y transmitirlos a la siguiente capa.
- **Capas ocultas:** Transforman los datos aplicando funciones de activación no lineales.
- **Capa de salida:** Genera la predicción final (Goodfellow et al., 2016).

El aprendizaje en las RNA se realiza mediante retropropagación del error, ajustando los pesos de las conexiones entre neuronas para minimizar el error, utilizando el método de optimización por descenso del gradiente (LeCun et al., 2015). Aunque las RNA son muy flexibles y capaces de modelar relaciones complejas, su interpretabilidad es limitada en comparación con algoritmos más sencillos.

4.5.3. Clasificación con Máquinas de Soporte Vectorial (SVM)

El uso de SVM para la clasificación de géneros musicales ha demostrado ser eficaz debido a su capacidad para manejar datos de alta dimensionalidad y encontrar la separación óptima entre las clases. SVM es particularmente útil cuando las clases no son linealmente separables, lo cual es común en muchos problemas de clasificación musical.

Máquina de Soporte Vectorial

Las SVM fueron introducidas por Vladimir Vapnik y su equipo en los años 90, y su principio fundamental es encontrar un *hiperplano* en un espacio de características de alta dimensionalidad que maximice el margen entre las clases de datos. El margen es la distancia entre el hiperplano y los puntos de datos más cercanos de cada clase, llamados *vectores de soporte*.

El objetivo es maximizar el margen para obtener un modelo que generalice mejor a datos no vistos. Este problema se formula como una optimización matemática:

Maximizar:

$$\frac{2}{\|W\|} \quad (4-8)$$

Sujeto a:

$$y_i(W \cdot X_i + b) \geq 1 \quad (4-9)$$

Donde: - W es el vector normal al hiperplano, - b es el sesgo, - X_i son los puntos de datos, - y_i son las etiquetas de las clases.

Kernel Trick: Para problemas no lineales, se utiliza el "kernel trick", que permite mapear los datos originales a un espacio de características de mayor dimensionalidad donde los datos se pueden separar linealmente. Los kernels comunes incluyen:

- **Kernel Lineal:** Para datos linealmente separables.
- **Kernel Polinómico:** Mapea los datos a un espacio de características polinómico.
- **Kernel Radial Basis Function (RBF):** También conocido como kernel gaussiano, es efectivo para datos no linealmente separables.
- **Kernel Sigmoide:** Similar a una red neuronal con una capa oculta.

El uso de Kernels en las máquinas de soporte vectorial permite que este algoritmo funcione aún cuando los datos no pueden separarse linealmente en su forma original. En el caso de clasificación de géneros musicales, las características espectrales que se extraen crean un espacio con una dimensión alta, en donde las divisiones entre géneros musicales no siempre son lineales. Por medio del Kernel trick, es posible transformar ese espacio en otro donde los distintos géneros puedan separarse usando hiperplanos. Esto permite que el modelo tenga mayor capacidad para aprender algunos patrones en la música, lo que ayuda en la clasificación. Es por esto que los kernels son claves para que las máquinas de soporte vectorial se adapten a la naturaleza no lineal de las señales musicales.

5. Metodología

5.1. Recolección de Datos

Para la caracterización espectral de los géneros musicales colombianos, Vallenato y Currulao, se inició la recolección de datos mediante la compilación de grabaciones en formato WAV de cada género. Se estudiaron 44 canciones del género tradicional Currulao y 45 canciones del género tradicional Vallenato, las canciones fueron descargadas de diversas fuentes disponibles en internet, asegurando que pertenecieran claramente a cada uno de los dos géneros analizados, basado en el género asignado por el autor de la canción al momento de su creación, las canciones tienen una duración promedio de entre 3 y 5 minutos. Posteriormente, los archivos fueron almacenados en directorios específicos en una unidad local, organizados por género para facilitar su acceso y manipulación.

5.2. Instalación y Configuración del Entorno

Se configuró el entorno de trabajo para garantizar la **reproducibilidad** y **funcionalidad del análisis**. Para ello, se instalaron paquetes esenciales del lenguaje de programación **R**, utilizados en la *manipulación de audio*, *análisis estadístico* y *aprendizaje automático*, tales como:

- **tuneR** para manipulación de señales de audio.
- **e1071** y **caret** para modelos estadísticos y aprendizaje automático.
- **ggplot2** y **corrplot** para visualización.
- **dplyr** y **reshape2** para transformaciones de datos.

El uso de estos paquetes, reconocidos por su aplicabilidad en el análisis de señales y el aprendizaje automático, permitió asegurar la compatibilidad entre procesos y la reproducibilidad de los resultados obtenidos.

5.3. Configuración de Directorios

Se estableció una estructura organizada de directorios con el fin de facilitar el acceso, clasificación y procesamiento de los datos.

En particular, se definieron dos directorios de entrada correspondientes a los géneros musicales analizados:

- **Currulaos:** Contiene archivos .wav para este género.
- **Vallenatos:** Contiene archivos .wav para este género.

5.4. Extracción de Características del Audio

Se extrajo un conjunto de características multivariadas representativas de las señales de audio, con el objetivo de capturar información relevante para su posterior análisis y clasificación.

Características extraídas:

- **Características temporales:** Energía, varianza, asimetría (skewness), concentración (kurtosis) y cruces por cero.
- **Características espectrales:** Centroides, ancho de banda y densidad espectral de potencia (calculada mediante el método de Welch).
- **Ciclos:** Frecuencia y duración de ciclos detectados en la señal.

Este conjunto de variables permite describir las propiedades de las señales en los dominios temporal y frecuencial, siendo aspectos fundamentales para la diferenciación entre géneros musicales.

5.5. Estimación de la densidad espectral

5.5.1. Procesamiento de Señales

Los datos de audio fueron transformados a un formato adecuado para su análisis estadístico y su aplicación en algoritmos de aprendizaje automático. Cada archivo fue procesado de manera individual, y se extrajeron de él diversas características relevantes, las cuales fueron almacenadas en un marco de datos (data frame). A cada observación se le asignó una etiqueta correspondiente al género musical de la canción. La normalización y organización de las características facilitó tanto el análisis comparativo como el entrenamiento de modelos de clasificación, asegurando la coherencia y consistencia del conjunto de datos.

5.5.2. Aplicación del Método de Welch

El método de Welch se empleó para estimar la densidad espectral de potencia (PSD) de las señales de audio, lo cual permitió:

- **Estimación de potencia en frecuencias específicas:** Facilitando la identificación de regiones frecuenciales características de cada género.
- **Extracción de características espectrales:** Como los centroides espectrales y el ancho de banda, útiles para los procesos de clasificación.
- **Comparación géneros musicales:** Las diferencias observadas en los periodogramas promedio proporcionaron criterios adicionales para la diferenciación entre géneros.

5.6. Análisis para identificar patrones y estructuras

Con el objetivo de identificar patrones y diferencias significativas entre los géneros musicales analizados, se llevó a cabo un análisis exploratorio de las características extraídas. Para ello, se emplearon gráficos de densidad y diagramas de caja (*boxplots*), lo que permitió visualizar y comparar la distribución de cada variable entre los géneros. Asimismo, se construyó una matriz de correlación, con el fin de detectar relaciones relevantes entre las distintas características. Este análisis preliminar no solo permitió validar la calidad de los datos obtenidos, sino también evaluar la capacidad discriminativa de las variables, lo que resulta fundamental para el posterior entrenamiento de los modelos de clasificación.

Comparación de Espectros Promedios

Con el propósito de analizar las diferencias espectrales entre los géneros musicales estudiados Vallenato y Currulao, se realizó una comparación de los espectros promedio de la potencia. Para ello, se utilizó el método de Welch, que permite estimar la densidad espectral de potencia de una señal mediante el promedio de periodogramas parciales. Este procedimiento se aplicó a cada género por separado, obteniendo así una representación promedio del contenido frecuencial de sus respectivas señales. Posteriormente, se realizó una comparación gráfica de estas densidades espectrales, lo que permitió identificar diferencias relevantes en la distribución de energía a lo largo del espectro de frecuencias. Estas diferencias reflejan patrones rítmicos y armónicos característicos de cada género, contribuyendo al entendimiento de sus particularidades acústicas y facilitando su clasificación automática.

5.7. Selección de características

Un paso fundamental en el procesamiento de señales de audio para la clasificación musical es la selección de características. Una selección adecuada permite extraer información representativa de cada género musical y reducir la dimensionalidad del conjunto de datos sin comprometer la relevancia de las variables utilizadas (Guyon and Elisseeff, 2003). En este estudio, se optó por un conjunto de características temporales y espectrales con el propósito de capturar diferencias distintivas entre los géneros Vallenato y Currulao.

Para el análisis de la dinámica rítmica, se consideraron características temporales como la energía total y la varianza de la señal, las cuales permiten describir la intensidad y la variabilidad del sonido (Tzanetakis and Cook, 2002). Por otro lado, para caracterizar el equilibrio frecuencial de cada género, se emplearon características espectrales como el centroide espectral y el ancho de banda, que permiten distinguir estilos con mayor presencia de frecuencias graves o agudas (Peeters, 2004).

La combinación de estos descriptores proporciona un enfoque robusto para la clasificación musical. A continuación, se detallan las características seleccionadas junto con una breve justificación de su uso, destacando su relevancia en la diferenciación de los géneros analizados.

Energía total: Esta característica es muy importante para distinguir la intensidad rítmica de los géneros. Los géneros como el Currulao presentan mayores niveles de energía al tener mayor percusión, a diferencia de géneros suaves o melódicos, como el Vallenato (Tzanetakis and Cook, 2002).

Varianza de la señal: La varianza refleja el grado de fluctuación en la amplitud de la señal a lo largo del tiempo, siendo un indicador clave de su dinamismo. En el caso del Currulao, caracterizado por ritmos marcados y golpes acentuados, se observa una varianza elevada. En contraste, el Vallenato, con una estructura más melódica y suave, presenta una menor variabilidad en su señal (Peeters, 2004).

Skewness (asimetría de la señal): La asimetría permite identificar cómo se distribuyen los picos de amplitud de la señal, en esta característica también influye los golpes musicales. En géneros con golpes rítmicos fuertes, como el Currulao, la asimetría es más pronunciada en comparación con géneros de estructura más homogénea como el Vallenato (Lartillot and Toivainen, 2007).

Kurtosis (curtosis de la señal): Evalúa la altura de los picos en la señal, lo que ayuda a diferenciar los ritmos percusivos marcados de los ritmos con melodías uniformes. Géneros con mayor golpe tienden a presentar valores de curtosis más altos (Peeters, 2004).

Cruces por cero: Relacionado con la rapidez de variabilidad de la señal, esta característica permite diferenciar estructuras rítmicas discontinuas de aquellas más continuas. Se ha descubierto que los géneros con líneas melódicas más sostenidas, como el Vallenato, suelen presentar mayor cruces por cero en comparación con los géneros como el Currulao más percusivos (Tzanetakis and Cook, 2002).

Centroid espectral: Representa el equilibrio tonal de las señales musicales. En géneros con predominancia de percusión y golpe, como el Currulao, el centroid espectral normalmente es más bajo, por otro lado en el Vallenato, donde predominan instrumentos melódicos como el acordeón, el valor del centroid espectral es más alto (Peeters, 2004).

Ancho de banda espectral: Es útil para medir la dispersión de las frecuencias en la señal. Se ha demostrado que géneros con una variedad de sonidos percusivos como el Currulao tienden a tener un mayor ancho de banda espectral, lo que facilita su diferenciación de géneros melódicos como el Vallenato (Lartillot and Toivainen, 2007).

5.8. Clasificación binaria de generos musicales

En este apartado se comparan tres métodos de clasificación con el objetivo de seleccionar el modelo que presenta el mejor desempeño en la tarea de clasificar los géneros musicales. Los algoritmos analizados son la Máquina de Vectores de Soporte (SVM), Random Forest (RF) y Redes Neuronales (NN). La comparación se realizó entre las medidas de exactitud (accuracy), Kappa, sensibilidad (recall) y especificidad.

Ahora bien, los datos utilizados fueron divididos en un conjunto de entrenamiento 70 % y un conjunto de prueba 30 %. Además, se normalizaron las variables para garantizar que la escala entre modelos fuera homogénea. Cada algoritmo fue entrenado y evaluado utilizando la matriz de confusión y las métricas correspondientes.

5.8.1. Clasificación mediante SVM

Este estudio contempla el desarrollo de un modelo de clasificación con la capacidad de diferenciar con precisión los géneros musicales Vallenato y Currulao, utilizando técnicas de aprendizaje automático aplicadas al análisis de señales de audio. Se plantea una estrategia basada en el uso de características espectrales y temporales extraídas directamente de las grabaciones, con el objetivo de entrenar un modelo eficiente que logre una alta tasa de acierto en la categorización de nuevas canciones.

En este contexto, se formulan dos hipótesis que guían el desarrollo de la investigación: en primer lugar, se plantea que es posible diferenciar los géneros Vallenato y Currulao a partir de características espectrales derivadas de las señales de audio; en segundo lugar, se postula que un modelo de clasificación entrenado con estas características puede alcanzar una precisión igual o superior al 75 % al predecir el género de nuevas canciones (Cortes and Vapnik, 1995). Estas hipótesis, aunque distintas, son complementarias en tanto que la validez de la segunda depende en gran medida de la solidez de la primera.

Para la implementación del modelo, se eligieron las Máquinas de Soporte Vectorial (SVM), ya que son particularmente eficaces en tareas de clasificación binaria con conjuntos de datos complejos y de alta dimensión (Cortes and Vapnik, 1995). El procedimiento realizado se detalla a continuación:

- División del conjunto de datos en subconjuntos de entrenamiento (70 %) y prueba (30 %).
- Normalización de las variables mediante centrado y escalado.
- Entrenamiento del modelo SVM utilizando un kernel radial.
- Evaluación del desempeño del modelo mediante matriz de confusión y métricas asociadas (precisión, sensibilidad, especificidad, entre otras).

Además del modelo SVM, se comparó su desempeño con dos algoritmos de clasificación, Random Forest y Redes Neuronales, con el fin de identificar cuál ofrece un mejor equilibrio entre precisión y capacidad de generalización. Finalmente, se evaluó la eficiencia del modelo seleccionado al aplicarlo en un escenario real, clasificando nuevas canciones de los géneros estudiados y determinando qué características del audio resultan más relevantes para la diferenciación automática de estilos musicales.

5.8.2. Clasificación de Nuevas Canciones

Una vez entrenado el modelo SVM, se procedió a aplicarlo sobre un conjunto de canciones no utilizadas previamente, con el fin de evaluar su capacidad para predecir correctamente

5 Metodología

el género musical. El procedimiento incluyó el procesamiento de las nuevas canciones para extraer las mismas características utilizadas durante la etapa de entrenamiento. Posteriormente, estas características fueron normalizadas utilizando los parámetros obtenidos del conjunto de entrenamiento, garantizando así la consistencia en la escala de los datos. En aquellos casos en que el modelo no logró asignar de manera confiable una clase conocida, se etiquetaron las canciones como “Desconocido”, lo cual permite identificar instancias que no pertenecen a los géneros previamente entrenados. Este proceso permitió evaluar la capacidad de generalización del modelo, así como su efectividad para reconocer patrones fuera del dominio aprendido.

6. Resultados

Con el fin de dar respuesta a la pregunta de investigación planteada, desarrollar los objetivos propuestos y realizar la implementación de la metodología descrita en el documento, se presenta esta sesión de resultados.

Inicialmente, se presenta una breve descripción de los datos utilizados para este estudio, y el análisis de las series temporales al utilizar el método de Welch y una vista a las características acústicas de los géneros musicales estudiados. Luego se presenta el análisis de densidades con gráficos de distribución de potencia en frecuencia mediante tablas descriptivas por cada género y su respectiva interpretación. Posteriormente, el resultado de la selección del algoritmo de Máquinas de Soporte Vectorial.

Para finalizar, dando cumplimiento con el objetivo de proponer una regla de clasificación para los ritmos musicales estudiados a partir de las densidades espectrales, se desarrolla la calificación automática de los géneros musicales con la implementación de un modelo de aprendizaje automático. Se muestra de la selección del mejor algoritmo mediante la comparación de métricas estadísticas como la exactitud, sensibilidad, especificidad, entre otras; esta comparación se realiza aplicando tres algoritmos a las características espectrales y temporales resultantes del análisis de densidades espectrales. Posterior a esto, se muestra la selección de características, con el fin de seleccionar las que más aporten al modelo. Luego, se muestra la experimentación realizada al algoritmo seleccionado para optimizar el algoritmo. Por último, se muestra el resultado de la clasificación automática mediante Maquinas de Soporte Vectorial.

6.1. Descripción de los Datos

Para esta investigación se seleccionaron 44 canciones del género tradicional Currulao y 45 canciones del género tradicional Vallenato, interpretadas por agrupaciones representativas de la región Pacífica y de la costa Atlántica del país, respectivamente. Ambos repertorios de canciones tienen una duración promedio de entre 3 y 5 minutos. Los archivos, en formato WAV, fueron descargados de internet para su análisis.

6.2. Análisis de Series Temporales

Al utilizar el método de Welch en esta investigación, los espectrogramas generados representan la densidad espectral de potencia, lo que asegura que los valores obtenidos sean promedios calculados en ventanas solapadas, proporcionando una representación suave y robusta del espectro. Este

6 Resultados

análisis permite identificar las principales características acústicas del Currulao, contribuyendo a su caracterización.

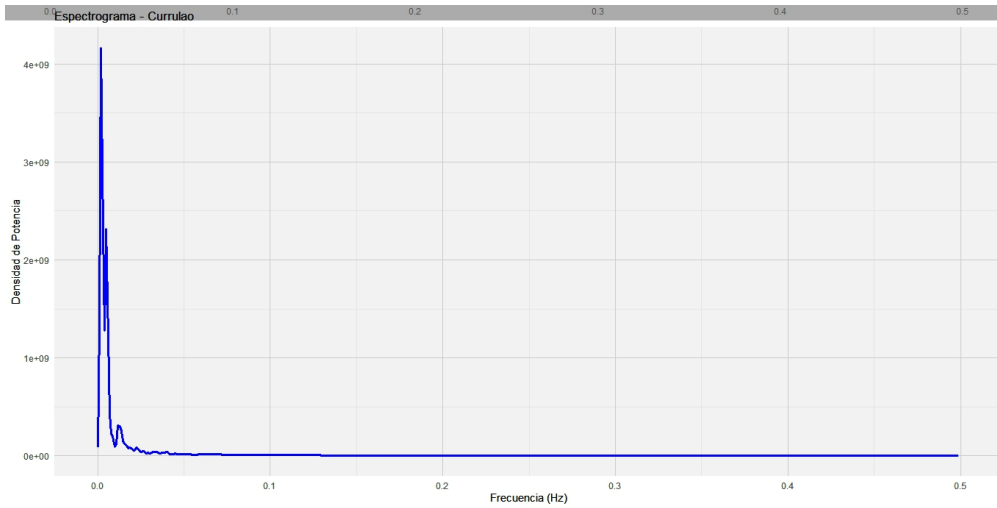


Figura 6-1.: Espectrograma del Currulao.

En la Figura **6-1** se muestra una alta concentración de energía en las frecuencias bajas las cuales están por debajo de $0,1Hz$, alcanzando picos superiores a 4×10^9 en densidad de potencia. Esto refleja el predominio de sonidos graves y ritmos marcados en este género, que suelen estar asociados con instrumentos como los tambores tradicionales y el bombo. A partir de $0,1Hz$, la densidad de potencia disminuye rápidamente, siendo prácticamente nula en frecuencias más altas específicamente las mayores a $0,2Hz$, lo que indica que las frecuencias altas no desempeñan un papel importante en la estructura acústica del Currulao.

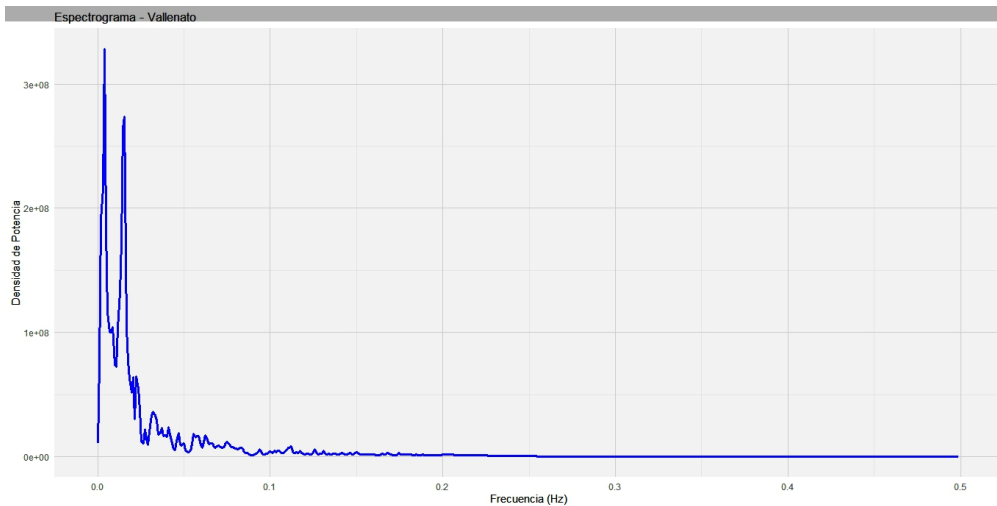


Figura 6-2.: Espectrograma del Vallenato.

Por otro lado, en la Figura **6-2** se muestra que en las frecuencias bajas, las cuales son $< 0,1Hz$, hay una alta densidad de potencia en este rango, con picos significativos que reflejan una gran cantidad

de energía concentrada en las frecuencias graves. Esto es característico de los instrumentos rítmicos del Vallenato, como la caja y la guacharaca, que generan patrones repetitivos con predominancia en el rango de frecuencias bajas (Luque, 2012).

En las frecuencias medias $0,1Hz - 0,2Hz$ se evidencia una disminución gradual en la densidad de potencia, pero se observan contribuciones relevantes, lo que refleja el uso de instrumentos melódicos como el acordeón. Este instrumento tiene un rango tonal más amplio que incorpora frecuencias medias, diferenciándolo de otros géneros como el Currulao (Luque, 2012).

Por último, en las frecuencias que son mayores a $0,2Hz$, la densidad de potencia cae rápidamente hasta valores cercanos a cero. Esto evidencia que las frecuencias altas no tienen un papel relevante en la estructura acústica del Vallenato, lo cual se alinea con su instrumentación tradicional.

Ahora bien, aunque ambos géneros presentan alta densidad de potencia en frecuencias bajas, el Vallenato destaca por su mayor diversidad tonal en el rango de frecuencias medias. Este comportamiento es consistente con la participación del acordeón como instrumento principal. Sin embargo, al comparar ambos géneros a través de un periodograma, podrían evidenciarse algunas similitudes y diferencias (Rey Bernal, 2019).

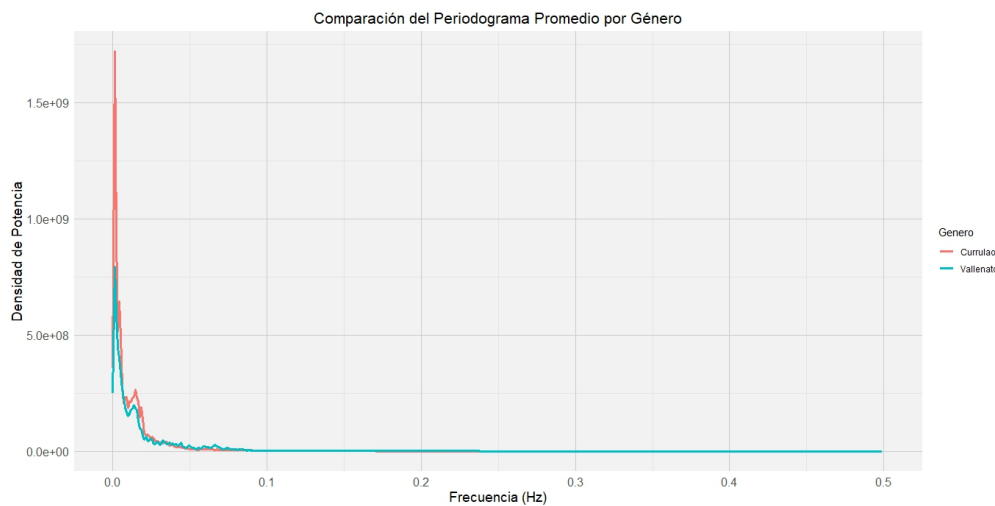


Figura 6-3.: Comparación del periodograma promedio por género.

En la Figura 6-6, se observa una comparación entre los periodogramas promedio del Currulao y el Vallenato, que revela tanto similitudes como diferencias relevantes entre estos géneros musicales. Ambos presentan alta densidad de potencia en frecuencias bajas por debajo de $0,1Hz$, lo cual es consistente con su carácter rítmico predominante (Rey Bernal, 2019). Por otro lado, al realizar el análisis a fondo, se destacan similitudes y diferencias esenciales que enriquecen la caracterización:

- Los dos géneros muestran su mayor densidad de potencia en sus frecuencias bajas, lo que puede indicar que los sonidos graves y las frecuencias fundamentales son muy importantes en sus estructuras musicales. Esto ayuda a reforzar el papel de los instrumentos de percusión en los dos géneros, como lo es el bombo en el Currulao y la caja en el Vallenato (Luque, 2012).
- El género del Currulao logra alcanzar picos significativamente más altos que el género del Vallenato en las frecuencias bajas. Esto refleja una mayor intensidad energética en los

elementos rítmicos del Currulao, probablemente atribuida a instrumentos como los cununos y el bombo, que generan patrones más marcados y de mayor potencia(Rey Bernal, 2019)

- A medida que la frecuencia aumenta, la densidad de potencia en el Currulao desciende rápidamente, mientras que en el Vallenato la transición es más gradual. Este patron en el Vallenato se podría asociar con la participación de algunos instrumentos melódicos como por ejemplo el acordeón, que logra aportar contenido en frecuencias medias, dando así un carácter tonal de más valor(Martínez, 2018).
- En los dos géneros musicales, las frecuencias que son mayores a 0.1 Hz tienen una densidad de potencia se podría interpretar como nula. Esto reafirma que los sonidos agudos no logran desempeñar un papel realmente relevante en la estructura acústica de estos dos géneros.

En la Figura **6-6** se muestra que, a pesar de que el Currulao y el Vallenato comparten una base acústica en frecuencias bajas, presentan algunas diferencias significativas en la intensidad de sus picos y en la distribución de la densidad de potencia en el espectro. Con este análisis se logra entender la estructura musical de los dos géneros y poder destacar sus particularidades desde un enfoque técnico y también cultural (Rey Bernal, 2019).

6.3. Análisis de densidades

La distribución de potencia describe cómo se concentra la energía acústica —es decir, la densidad de potencia en torno a una frecuencia determinada, teniendo en cuenta los diferentes valores de amplitud. En términos prácticos, esta distribución permite observar cómo varía la densidad de potencia a lo largo de un rango de valores y facilita la detección de patrones característicos en el comportamiento energético de la señal.

En este estudio, el análisis se enfoca en la frecuencia 10 para ambos géneros musicales. Cada punto del gráfico correspondiente representa la probabilidad de que la densidad de potencia, en una muestra de Currulao o Vallenato, adopte un valor específico. Con base en estos datos, se presenta a continuación la distribución de potencia observada en la frecuencia 10 para los dos géneros analizados.

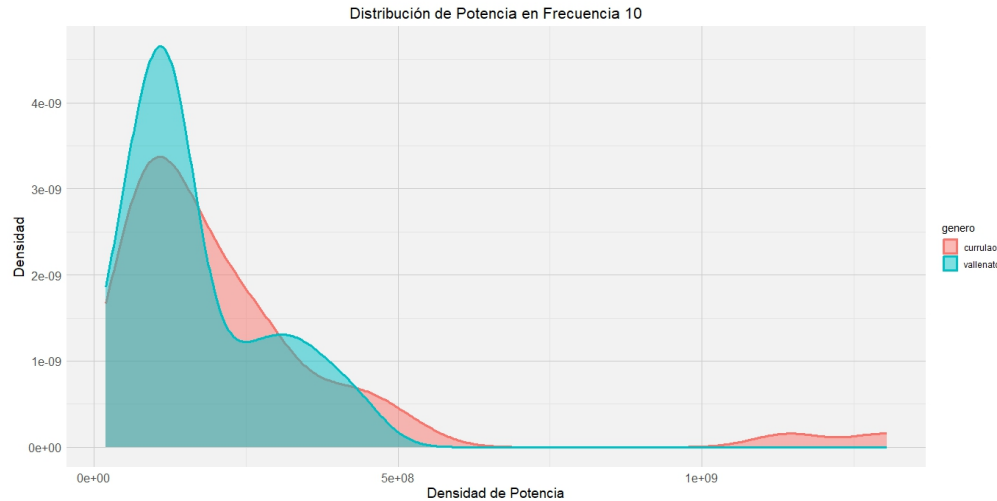


Figura 6-4.: Distribución de potencia en frecuencia 10.

En la Figura 6-4, se muestra cómo es la distribución de potencia en la frecuencia 10 para el Currulao y cómo lo es para el Vallenato. En esta gráfica se muestra cómo se distribuyen las densidades de potencia dentro de cada uno de los géneros, permitiendo así observar similitudes o diferencias clave. En esta figura se presenta la frecuencia 10, porque con esta frecuencia se permite ilustrar la variabilidad de las señales en una dimensión específica.

- El género de Vallenato presenta un pico más alto en la densidad de potencia, lo que puede indicar una mayor energía promedio en esta frecuencia particular. Esto se podría asociar con predominancia del acordeón en este ritmo musical, que genera señales con mayor amplitud en este rango de frecuencia (Caraballo, 2024).
- Por otro lado, el Currulao muestra un pico más bajo, reflejando un comportamiento más moderado en esta frecuencia, posiblemente relacionado con el predominio de los tambores y otros instrumentos percusivos que generan patrones rítmicos de menor amplitud.
- Existe un rango donde las densidades de ambos géneros se solapan, indicando que, en ciertos valores de potencia, las características acústicas de ambos géneros son similares.
- En valores de potencia más altos (por encima de 1×10^9), el Currulao muestra una pequeña contribución, mientras que el Vallenato es prácticamente inexistente en este rango.

La Figura 6-4 aporta información clave sobre el comportamiento energético del Currulao y el Vallenato en la frecuencia 10, de este se concluye que el Vallenato presenta una mayor densidad de potencia promedio, probablemente vinculada a su instrumentación melódica y que el Currulao, aunque tiene menor densidad promedio, muestra contribuciones en rangos de mayor potencia, destacando su carácter rítmico (Rey Bernal, 2019).

Es importante tener en cuenta el análisis de la densidad de energía por cada uno de los géneros. Ya que esta describe cómo se dispersan los valores de energía de las señales dentro de un rango específico. En este caso, representa la intensidad promedio del sonido producido por los instrumentos y cómo

6 Resultados

esta varía entre las señales de cada género y es útil para identificar patrones característicos que diferencian los géneros musicales, especialmente en términos de intensidad y variabilidad (Martínez, 2018).

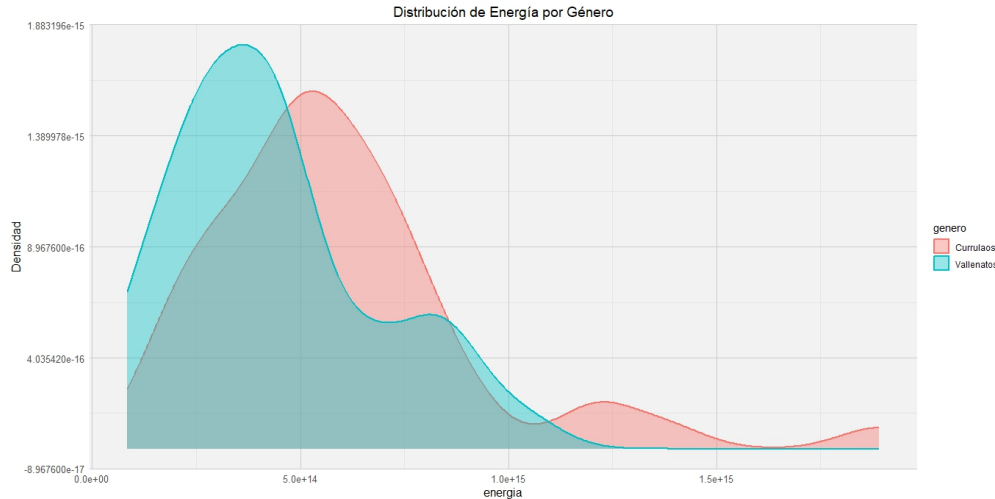


Figura 6-5.: Densidad de energía por género.

En la Figura 6-5, se muestra como se distribuye la energía para los géneros Currulao y Vallenato. Este gráfico representa la distribución de energía dentro de cada género, dejando ver similitudes y diferencias importantes en su comportamiento acústico:

- El Currulao alcanza un pico de densidad de energía más alto alrededor de 5×10^{14} , lo que indica que este rango es el más representativo de la energía de las señales de este género. Este comportamiento podría estar relacionado con la alta intensidad rítmica generada por los tambores y otros instrumentos de percusión dominantes en el Currulao (Cruz et al., 2019).
- El Vallenato, por su parte, presenta un pico más bajo pero más amplio, ubicado alrededor de 2×10^{14} . Esto refleja una menor intensidad en términos de energía promedio, pero con mayor diversidad en los patrones de energía, lo que podría atribuirse al carácter melódico del acordeón combinado con los elementos rítmicos (UNESCO, 2015).
- Existe una región considerable de superposición entre las distribuciones de ambos géneros, especialmente en el rango de 2×10^{14} a 8×10^{14} . Esto indica que, en este rango de energía, los géneros comparten características similares, lo que puede dificultar su diferenciación (Mendes et al., 2010).
- En los valores de energía (por encima de $1,5 \times 10^{15}$), el Currulao presenta contribuciones relevantes, mientras que el Vallenato tiene una presencia casi nula. Esto sugiere que algunas señales del Currulao contienen episodios de mayor intensidad energética la cual posiblemente este ligada a fenómenos rítmicos como la polirritmia afrodescendiente, probablemente asociados con patrones rítmicos más marcados.

La Figura 6-5 aporta información clave sobre la distribución de energía en los géneros musicales Currulao y Vallenato, como que el Currulao presenta picos de energía más altos, relacionados con

su carácter rítmico, mientras que el Vallenato tiene una distribución más amplia y moderada, lo que refleja su enfoque melódico. También que en las áreas de superposición entre los géneros sugieren similitudes que podrían complicar su clasificación únicamente basándose en la energía (Cruz et al., 2019).

Este análisis resalta la importancia de combinar características adicionales para una diferenciación más precisa entre los géneros, por ello en el siguiente apartado se van analizar otras características para caracterizar ambos géneros musicales.

6.4. Características generales y espectrales

Características Generales del Género Currulao

Antes de analizar las características extraídas, se presenta a continuación una tabla con un conjunto de señales del género Currulao, procesadas para ilustrar los valores típicos de las métricas utilizadas en el estudio. Esta selección permite observar cómo se comportan las principales características estadísticas y espectrales asociadas a este género.

Tabla 6-1.: Características generales de las señales del género Currulao.

Energía	Varianza	Simetría	Kurtosis	Cruces Cero	Centroide	A.banda
1.1984e+15	87149226	0.0062	0.4624	1	0.0128	0.0261
1.1443e+14	32453100	-0.0217	1.4595	0	0.0104	0.0229
3.7215e+14	33695462	0.0522	1.2716	0	0.0276	0.0425
7.6152e+14	49197605	0.0059	1.2051	0	0.0116	0.0195
2.7775e+14	18209153	0.0256	1.8609	0	0.0069	0.0139
6.5345e+14	35951447	0.0168	1.8850	0	0.0107	0.0173

La Tabla **6-1** muestra las características generales estudiadas de las señales del Currulao. Estas métricas permiten analizar aspectos clave de la estructura acústica del género, proporcionando información sobre su comportamiento energético, espectral y rítmico:

- Los valores de energía se encuentran en el rango de 10^{14} a 10^{15} , con algunos valores más altos como $1,1984 \times 10^{15}$. Lo cual indica que el género de Currulao tiene un carácter energético alto, evidenciando la predominancia de los instrumentos de percusión que pueden generar patrones rítmicos fuertes y constantes (Cruz et al., 2019; Toledo et al., 2019). Este nivel de energía se muestra como es un rasgo distintivo del género Currulao y sugiere una señal intensa y bastante definida.
- La varianza oscila entre $1,82 \times 10^7$ y $8,71 \times 10^7$. Esto expresa de una forma clara la dispersión de las señales, indicando así que hay cambios en las amplitudes, y se puede ver relacionada con los comportamientos dinámicos presentes en las composiciones del Currulao (Lartillot and Toivainen, 2008). Esto refuerza su carácter percusivo y rítmicamente diverso.

6 Resultados

- Los valores de simetría están mayormente cercanos a cero (positivos o negativos), con un rango de -0.0217 a 0.0522. Esto indica que las señales del Currulao son simétricas en su distribución de amplitud, pero presentan ligeras asimetrías que podrían deberse a la naturaleza irregular de los golpes de percusión (Peeters, 2004).
- Los valores de kurtosis varían entre 0.46 y 1.86, lo que indica distribuciones con colas ligeras y picos moderados. Esto significa que las señales del Currulao tienden a tener patrones relativamente uniformes, sin cambios abruptos en amplitud (Tzanetakis and Cook, 2002). Este comportamiento es consistente con un género que se basa en ritmos repetitivos y consistentes.
- El valor predominante de 0 en los cruces por cero sugiere que las señales del Currulao tienen pocas oscilaciones alrededor del eje cero. Esto es típico de señales dominadas por bajas frecuencias, donde los golpes rítmicos generan amplitudes que permanecen mayormente en un solo lado del eje (Jensen et al., 2006).
- Los valores bajos del centroide (rango de 0.0069 a 0.0276) y del ancho de banda (0.0139 a 0.0425) indican que las frecuencias bajas son predominantes en el Currulao (Tzanetakis and Cook, 2002). Esto refuerza el carácter rítmico y grave del género, donde los instrumentos de percusión juegan un papel fundamental.

La Tabla **6-1** revela que las señales del Currulao que se tomaron para este análisis están caracterizadas por los altos niveles de energía, consistentes con su carácter rítmico e intenso, la predominancia de frecuencias bajas, reflejando el papel principal de la percusión y la baja variabilidad en los patrones, lo que refuerza su estructura repetitiva y rítmica (Cruz et al., 2019; Toledo et al., 2019).

Frecuencias Espectrales del Género Currulao

Tabla 6-2.: Frecuencias específicas de las señales del género Currulao.

Frec 10	Frec 20	Frec 30	Frec 40	Frec 50	Frec 60	Frec 70	Frec 90
22915718	7384455	39242859	20162015	11017446	17552493	8177007	9428599
27370622	5611060	20403437	12816771	6065953	14650397	3727322	5338725
21055588	10389485	21357487	14269156	10121458	4800674	11406051	8459347
12588307	34887534	42936174	19206124	8506072	7731081	5421801	6134768
20343580	20602421	4370765	5017817	1851446	9285874	612267	3680882
7190737	4753267	20014692	6716827	3669794	2066088	2613690	1234846

La Tabla **6-2** presenta las frecuencias específicas de las señales del género Currulao. Estas frecuencias corresponden a valores representativos en diferentes bandas (10, 20, 30, ..., 90 Hz) y reflejan la intensidad energética asociada a cada rango. Este análisis permite identificar patrones rítmicos y armónicos relevantes para el género.

- Los valores en la frecuencia 10 son los más altos en comparación con otras bandas, alcanzando picos como $2,73 \times 10^7$. Esto indica que el Currulao concentra gran parte de su energía en las frecuencias más bajas, reforzando su carácter rítmico basado en instrumentos percusivos graves como el bombo y los cununos (Benetos et al., 2018; Benadon, 2006).
- En la frecuencia 30 se observa un aumento significativo de la energía, con valores que alcanzan hasta $4,29 \times 10^7$. Este comportamiento sugiere que las frecuencias bajas y medias también tienen relevancia en el Currulao, posiblemente relacionadas con las resonancias y patrones secundarios generados por los instrumentos (Scheirer, 1998; Tzanetakis and Cook, 2002).

La Tabla 6-2 evidencia que el Currulao tiene un perfil acústico dominado por frecuencias bajas (10 y 30 Hz), con una menor contribución de frecuencias medias y una energía casi nula en frecuencias altas. Este análisis refuerza el carácter rítmico y grave del género, vinculado a sus instrumentos de percusión tradicionales (Benetos et al., 2018; Tzanetakis and Cook, 2002).

Características Generales del Género Vallenato

Tabla 6-3.: Características generales de las señales del género Vallenato.

Energía	Varianza	Simetría	Kurtosis	Cruces Cero	Centroide	A.banda
2.7088e+14	20849553	0.2289	2.8633	0	0.0313	0.0446
8.4375e+14	6518795	0.0877	2.2633	0	0.0253	0.0362
8.3924e+14	63311842	-0.0604	1.1679	0	0.0236	0.0372
5.2832e+14	29586910	0.0577	2.2927	0	0.0480	0.0487
4.5659e+14	57706434	-0.1678	1.2082	0	0.0248	0.0358
2.0442e+14	15185430	0.0508	1.5518	0	0.0254	0.0513

La Tabla 6-3 muestra las características generales de las señales del género Vallenato. Estas métricas permiten comprender la estructura energética, espectral y rítmica de este género, destacando sus diferencias y similitudes con otros géneros musicales como el Currulao.

- Los valores de energía oscilan entre $2,04 \times 10^{14}$ y $8,43 \times 10^{14}$, con una variabilidad significativa. Esto indica que el Vallenato tiene señales de amplitud moderada a alta, probablemente debido a la combinación de instrumentos melódicos como el acordeón y rítmicos como la caja y la guacharaca (Benadon, 2006; Dixon, 2007). La amplitud más alta refleja momentos de mayor intensidad melódica, mientras que los valores más bajos podrían estar asociados a pasajes menos dinámicos.
- Los valores de varianza van desde $6,51 \times 10^6$ hasta $6,33 \times 10^7$. Esta dispersión evidencia una mayor diversidad en la intensidad de las señales, probablemente derivada de la interacción entre los elementos melódicos y rítmicos del género (Tzanetakis and Cook, 2002). Esto sugiere que las señales del Vallenato son más dinámicas en comparación con géneros puramente rítmicos como el Currulao.

6 Resultados

- Los valores de simetría varían entre -0.1678 y 0.2289, con la mayoría cercanos a cero. Esto indica que las señales tienen una distribución balanceada en amplitudes, simetrías en algunos casos. Estas simetrías podrían deberse a la predominancia de patrones melódicos, que tienden a generar variaciones más marcadas en la señal (Lartillot and Toivainen, 2008; Benetos et al., 2018).
- La kurtosis oscila entre 1.16 y 2.86, lo que sugiere que las señales consideradas para el género del Vallenato tienen distribuciones más concentradas y picos moderados en comparación con géneros como el Currulao. Este comportamiento está alineado con un género melódico donde los instrumentos generan picos regulares pero menos extremos (Scheirer, 1998; Jensen et al., 2006).
- El valor predominante de 0 en los cruces por cero indica que las señales consideradas tienen pocas oscilaciones alrededor del eje cero, lo que refleja su enfoque melódico y rítmico con predominancia de frecuencias fundamentales claras y armónicas bien definidas (Benadon, 2006).
- Para las señales consideradas los valores del centroide van desde 0.0236 a 0.0480, y los del ancho de banda desde 0.0358 a 0.0513. Esto indica que las señales consideradas tienen un contenido más notable en frecuencias medias en comparación con el Currulao. Estas métricas refuerzan el carácter melódico del Vallenato, donde instrumentos como el acordeón ocupan un rango más amplio en el espectro sonoro. (Benetos et al., 2018; Peeters, 2004).

La Tabla **6-3** muestra que las señales del Vallenato tienen: Una energía moderada a alta, con una variabilidad significativa que refleja su complejidad melódica y rítmica. Frecuencias medias más prominentes que el Currulao, reforzando su carácter melódico. Señales balanceadas pero dinámicas, derivadas de la interacción entre elementos melódicos y rítmicos.

Frecuencias Espectrales del Género Vallenato

Tabla 6-4.: Frecuencias específicas de las señales del género Vallenato.

Frec 10	Frec 20	Frec 30	Frec 40	Frec 50	Frec 60	Frec 70	Frec 90
10027864	7630130	22220231	23052981	19093507	10513691	10498222	3027480
24066591	20802559	6893523	7149628	2579530	995788	3304588	345523
18028910	31055365	2951887	6764757	4040037	3857132	2480479	1144379
19735381	96515300	50971309	5794873	2889358	3875172	2456047	927469
69583544	31056190	50971309	6740577	3105612	2758224	927469	335761

En la Tabla **6-4**, se observa que en las frecuencias 30 y 40 tienen ciertos valores particularmente altos, lo que sugiere que estas frecuencias son dominantes en el género Vallenato; en cambio, en las frecuencias como la frecuencia 90, posee valores bajos respecto al resto de frecuencias, lo que evidencia que esta frecuencia tiene poca influencia en el género Vallenato.

6.5. Clasificación

6.5.1. Comparación de tres algoritmos

A continuación, se presentan los resultados obtenidos para los algoritmos evaluados para selección del mejor modelo a desarrollar:

Tabla 6-5.: Comparación de Modelos de Clasificación para ambos géneros musicales

Modelo	Exactitud	Kappa	Sensibilidad	Especificidad
SVM	0.8077	0.6154	0.9231	0.6923
Random Forest	0.8462	0.6923	0.8462	0.8462
Redes Neuronales	0.5000	0.0000	1.0000	0.0000

En la tabla **6-5**, se puede observar que el modelo de Random Forest podría ser la mejor opción debido a su mayor exactitud y equilibrio entre sensibilidad y especificidad. A diferencia de la red neuronal, que presenta un sesgo en su clasificación, y del SVM, que tiene una menor especificidad, Random Forest demuestra un rendimiento más estable y confiable en la tarea de clasificación. sin embargo, existen razones estadísticas que justifican el uso de Maquinas de soportes vectoriales en lugar de Random Forest.

6.5.2. Mejor algoritmo

Dado que existen razones estadísticas que justifican el uso de Maquinas de soportes vectoriales en lugar de random Forest, vamos a mencionar algunas de ellas.

- Las maquinas de soporte vectorial son conocidas por su capacidad para generalizar bien en espacios de alta dimensión, especialmente cuando el número de muestras es limitado (Vapnik, 1998). Dado que el modelo de maquinas se basa en encontrar el hiperplano óptimo de separación entre clases utilizando los vectores de soporte más representativos, es menos susceptible a sobreajustarse al ruido del conjunto de entrenamiento (Cortes and Vapnik, 1995). En contraste, aunque RF reduce el sobreajuste mediante la combinación de múltiples árboles, su desempeño puede degradarse cuando hay muchas variables no informativas (Breiman, 2001).
- Si bien Random Forest ofrece una exactitud ligeramente superior, se observa que el modelo de SVM tiene una mejor sensibilidad 92,31 % en comparación con RF 84,62 %, lo que indica que SVM es más eficiente en identificar la clase Currulao. En problemas de clasificación donde uno de los géneros puede ser más difícil de reconocer (clase minoritaria), maximizar la sensibilidad es clave para evitar sesgos en la predicción (Rehan Akbani, 2004).
- El modelo de RF depende de la aleatoriedad en la selección de subconjuntos de datos y características en cada árbol, lo que puede generar variabilidad en los resultados

entre ejecuciones distintas (Louppe, 2014). Por otro lado, SVM resuelve un problema de optimización convexa, lo que garantiza que siempre encontrará la misma solución óptima para los mismos datos (Cristianini and Shawe-Taylor, 2000). Esta estabilidad es crucial cuando se requiere un modelo confiable y reproducible.

- Desde un punto de vista estadístico, el SVM se basa en maximizar el margen entre clases, lo cual tiene un fundamento sólido en la teoría de la generalización y minimización del error estructural (Structural Risk Minimization) (Vapnik, 1995). En contraste, RF, al ser un método basado en múltiples árboles de decisión, tiene una mayor complejidad y menor interpretabilidad (Trevor Hastie, 2009).

6.5.3. Selección de características

Se seleccionaron estas características porque expresan propiedades distintivas de cada género musical, facilitando la clasificación mediante un algoritmo de aprendizaje automático. Su utilización conjunta permite un análisis integral de los aspectos rítmicos y tonales, aportando de forma positiva a la precisión del modelo de clasificación. A continuación se presenta una tabla con las conclusiones extraídas de las anteriores descripciones.

Característica	Diferencia clave entre géneros
Energía	Más alta en Currulao por golpes rítmicos
Varianza	Más alta en Currulao por cambios de amplitud
Simetría	Más alto en Currulao por golpes de tambor
Curtosis	Más alto en Currulao por picos fuertes
Cruces por cero	Más alto en Vallenato por continuidad melódica
Centroid espectral	Más alto en Vallenato por notas agudas
Ancho de banda	Más alto en Currulao por diversidad de sonidos

Tabla 6-6.: Diferencias clave entre las características de Currulao y Vallenato

6.5.4. Experimentación con maquinas de soporte vectorial

Con el propósito de determinar las características con mayor relevancia en la clasificación de géneros musicales por medio de Maquinas de soportes vectoriales se realizó una importancia de las características para identificar cual tenía mayor peso en la predicción. El siguiente gráfico es el resultado.

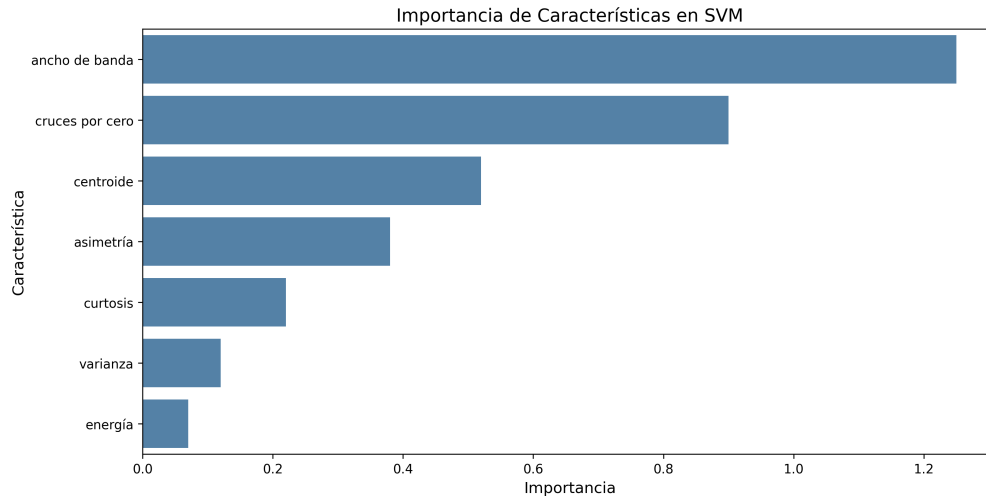


Figura 6-6.: Importancia de las características en las maquinas de soporte vectoriales.

En la gráfica **6-6** se observa que características como el Ancho de banda y el Centroide son las más relevantes para la clasificación de los géneros musicales. En contraste, variables como la varianza y la energía presentan menor impacto.

A partir de esta información, se diseñaron tres modelos de clasificación utilizando máquinas de vectores de soporte (SVM), en los cuales se varió la cantidad de variables incluidas. La selección de características se basó en su nivel de importancia, destacando las siguientes: ancho de banda, cruces por cero, centroide, simetría, curtosis, varianza y energía.

Los modelos propuestos fueron los siguientes:

- **Primer Modelo:** Este modelo contiene todas las características anteriormente mencionadas, las cuales son: Ancho de banda, Cruces cero, Centroide, Simetría, Curtosis, Varianza y Energía.
- **Segundo Modelo:** En este modelo se escogieron las mejores 5 características mencionadas en la importancia, las cuales fueron: Ancho de banda, Cruces cero, Centroide, Simetría, Curtosis.
- **Tercer Modelo:** En este último modelo se tuvieron en cuenta las 3 mejores características predichas por la importancia las cuales son: Ancho de banda, cruces cero, centroide.

Se plantearon las siguientes hipótesis:

Hipótesis

- H_1 : No hay diferencias significativas en la precisión entre los modelos SVM evaluados.
- H_2 : Al menos un modelo SVM tiene una precisión significativamente mayor que los otros.

Los resultados fueron los siguientes:

Modelo	Proporción de precisión
Primer Modelo	0.8077
Segundo Modelo	0.7692
Tercer Modelo	0.6923

Tabla 6-7.: Comparación de precisión entre modelos SVM

En la tabla **6-7**, se puede observar que el modelo con una mejor precisión para predecir es el primer modelo en donde están presente todas las características.

Para determinar si existen diferencias significativas en la precisión de los tres modelos SVM desarrollados, se aplicó una prueba de igualdad de proporciones, la cual permite comparar si las proporciones de éxito (precisión) entre los modelos son estadísticamente distintas. Esta prueba es útil cuando se desea evaluar el rendimiento de varios clasificadores sobre una misma tarea, bajo el supuesto de que los resultados siguen una distribución binomial (Dietterich, 1998).

Los resultados de esta prueba fueron los siguientes:

Estadístico	Valor
X^2 (Chi-cuadrado)	0.9741
Grados de libertad	2
Valor p	0.6144

Tabla 6-8.: Resultados de la prueba de igualdad de proporciones

En la Tabla **6-8**, se observa que el $valor - p$ es de 0.6144, teniendo en cuenta que tomamos como nuestro nivel de significancia $\alpha = 0,05$ no se puede rechazar la hipótesis Nula de que H_1 : No hay diferencias significativas en la precisión entre los modelos SVM evaluados. Lo anterior indica que no existe evidencia estadística para afirmar que existen diferencias significativas en la precisión de los 3 modelos.

Ahora bien, en la Tabla **6-7** se puede observar que la proporción de precisión va de manera descendente siendo de la siguiente manera: Primer Modelo: (0,8077), Segundo Modelo: (0,7692), Tercer Modelo: (0,6923) de acuerdo al test de igualdad de proporciones, los modelos no tienen diferencias significativas lo que desde el punto de vista estadístico ningún modelo supera en precisión al otro. Sin embargo, revisemos algunas razones para guiarnos del primer modelo con todas las características y utilizarlo en nuestra clasificación:

- Cuando múltiples modelos presentan diferencias no significativas en su desempeño, el criterio práctico es seleccionar aquel con la mejor métrica de precisión observada, dado que pequeñas mejoras pueden acumularse en escenarios de producción (Bishop, 2006).

- En ausencia de diferencias estadísticamente significativas, seleccionar el modelo con la mayor precisión empírica es una estrategia razonable para minimizar el error de predicción en futuras observaciones (Trevor Hastie, 2009).
- La falta de significancia estadística en las diferencias entre modelos no descarta la existencia de una diferencia real. En tales casos, elegir el modelo con mejor rendimiento observacional suele ser una decisión pragmática (Kuhn and Johnson, 2013).

Por las anteriores razones mencionadas para la realización de nuestro modelo de clasificación vamos a tomar en cuenta todas las características que se tuvieron en cuenta en el Primer Modelo, que fueron:

Ancho de banda, cruces cero, centroide, simetría, curtosis, varianza y energía.

6.5.5. Clasificación por medio de maquina de soporte vectorial

En el anterior apartado, se llegó a la conclusión de utilizar las siguientes características para nuestro modelo de clasificación: Ancho de banda, cruces cero, centroide, simetría, kurtosis, varianza y energía.

Al aplicar nuestro modelo de Maquinas de Soporte Vectorial para clasificar nuestros géneros musicales teniendo en cuenta las características anteriormente mencionadas, los siguientes fueron los resultados de la clasificación.

		Observación	
		Currulao	Vallenato
Predicción	Currulao	12	4
	Vallenato	1	9

Tabla 6-9.: Matriz de confusión

En la tabla **6-9**, se observa los resultados de la clasificación para los dos géneros musicales en donde se puede observar que:

- Se obtuvieron 12 canciones de Currulao correctamente clasificadas como Currulao.
- Resultaron 4 canciones de Vallenato clasificadas incorrectamente como Currulao.
- Hubo 1 canción de Currulao clasificada incorrectamente como Vallenato.
- Fueron 9 canciones de Vallenato correctamente clasificadas como Vallenato.

Métrica	Valor
Precisión (Accuracy)	80.77 %
Intervalo de Confianza (95 %)	(60.65 %, 93.45 %)
No Information Rate (NIR)	50.00 %
P-Valor (Acc > NIR)	0.001247
Kappa	0.6154
Sensibilidad	92.31 %
Especificidad	69.23 %
Valor Predictivo Positivo (VPP)	75.00 %
Valor Predictivo Negativo (VPN)	90.00 %
Balanced Accuracy	80.77 %
McNemar's Test P-Value	0.371093

Tabla 6-10.: Métricas de Evaluación del Modelo

En la tabla **6-10** se observan las métricas del modelo que nos indican que tan significativo es, de lo cual podemos decir lo siguiente:

- El modelo tuvo una precisión del 80,77, esta es la misma precisión obtenida en el apartado anterior debido a que fueron los mismos datos y las mismas características, por otro lado este resultado nos indica el modelo clasifica de manera adecuada el 80,77 % de las muestras.
- El rango de precisión en el cual se espera se encuentren clasificadas el 95 % de las muestras es (60,65 %, 93,45 %)
- El No Information Rate (NIR) de un 50 % indica que mide la precisión que se tendría si siempre se predijera la clase más frecuente en los datos, esta precisión sería del 50 %
- P-Valor (Acc > NIR) = 0.001247 Este valor es menor de 0.05, se puede decir que el modelo es mejor que una clasificación aleatoria.
- El valor Kappa fue de 0.6154 lo cual indica que existe una similitud entre las predicciones y los valores reales.
- La sensibilidad fue de 92,31 % lo que significa que en ese porcentaje nuestro modelo logra identificar el Currulao.
- La especificidad es de 69,23 % lo que indica que el modelo es capaz de identificar el Vallenato en ese porcentaje.
- Como el Valor Predictivo Positivo (VPP) es de 75,00 % si nuestro modelo predice Currulao, existe el 75 % de probabilidad que si lo sea.

- El Valor Predictivo Negativo (VPN) fue de 90,00 % esto quiere decir que si el modelo predice Vallenato, existe un 90 % de probabilidad que de si lo sea.
- Dado que el Balanced Accuracy es de 80,77 % este es un promedio entre la sensibilidad y la especificidad.
- En la prueba McNemar's Test el *ValorP* fue de 0.371093 lo que indica que no hay diferencias estadísticamente significativas en la clasificación de los errores.

6.5.6. Implementación de clasificación a nuevas canciones

Se realiza la calificación utilizando el mejor modelo desarrollado mediante el algoritmo de máquina de soporte vectorial. La evaluación se realiza con ocho canciones, cuatro correspondientes al género Vallenato y cuatro al género Currulao.

No.	Género real	Clasificación del modelo	Aciertos
1	Currulao	Vallenatos	✗
2	Currulao	Currulaos	✓
3	Currulao	Currulaos	✓
4	Currulao	Currulaos	✓
5	Vallenato	Vallenatos	✓
6	Vallenato	Vallenatos	✓
7	Vallenato	Vallenatos	✓
8	Vallenato	Vallenatos	✓

Tabla 6-11.: Clasificación de canciones

En la Tabla **6-11** se observa que la clasificación utilizando MSV muestra un desempeño mayormente acertado en la identificación de los géneros musicales, lo que indica un rendimiento aceptable. En general, el modelo presenta una precisión alta, clasificando correctamente 7 canciones, lo que corresponde a un 87 % de precisión global. En detalle, la precisión alcanza el 100 % para el género Vallenato y el 75 % para el Currulao.

7. Conclusiones y Recomendaciones

7.1. Conclusiones

El análisis de los espectrogramas y de las densidades espectrales de potencia permitió evidenciar características distintivas entre los géneros musicales Currulao y Vallenato. Ambos presentaron una alta concentración de energía en las frecuencias bajas; sin embargo, el Currulao mostró picos de potencia más elevados en comparación con el Vallenato, lo que sugiere una mayor intensidad en sus componentes rítmicos de baja frecuencia. En contraste, en las frecuencias altas, la potencia fue prácticamente nula para ambos géneros, lo que indica una clara predominancia de sonidos graves en sus estructuras sonoras. También se observan algunas superposiciones en la distribución espectral, lo que puede sugerir que hay similitudes, las cuales pueden dificultar, en algunos casos, la posterior clasificación automática entre los géneros.

La aplicación del método de Welch resultó adecuada para representar las propiedades frecuenciales de las señales, aportando información relevante para el análisis estadístico y mejorando la capacidad discriminativa en la clasificación de géneros musicales.

Al realizar el análisis de las características generales y frecuenciales de ambos géneros, se encontró que el Currulao presentó menor variabilidad en sus atributos como en la energía, la simetría y la curtosis. Esto podría interpretarse como una mayor consistencia en las señales musicales de este género. Por otro lado, el Vallenato mostró una mayor variabilidad y amplitud en estas características, lo cual puede reflejar una estructura rítmica más compleja y variable. En cuanto a las frecuencias dominantes, se evidenció que en el Currulao predominan las bandas bajas y medias, mientras que en el Vallenato se observan picos más pronunciados en las frecuencias medias, lo que coincide con los instrumentos melódicos que predominan como el acordeón.

Por último, el modelo de clasificación basado en Máquinas de Soporte Vectorial (SVM) mostró un desempeño sólido, alcanzando una precisión del 87,5 %. Este resultado muestra equilibrio entre sensibilidad y especificidad, indicando que el modelo fue capaz de identificar correctamente muestras de ambos géneros. Aunque se observó una mayor precisión en la clasificación del Vallenato, la identificación de las canciones de Currulao también fue satisfactoria. Estos resultados apoyan la hipótesis del estudio, al demostrar que existen diferencias acústicas suficientes entre ambos géneros para permitir su clasificación automática mediante técnicas de aprendizaje automático.

7.2. Recomendaciones

Basado en los resultados obtenidos, se proponen varios puntos de mejora para futuras investigaciones de caracterización espectral y en clasificación automática de géneros musicales. En primer lugar, se sugiere explorar métodos alternativos de caracterización espectral, como el Análisis Singular Espectral (SSA, por sus siglas en inglés), que podría proporcionar una representación más precisa de las componentes frecuenciales de la señal. Asimismo, se recomienda ampliar el conjunto de datos a analizar, incorporando un mayor número de canciones por género, por ejemplo, quinientas canciones de cada género musical, con el fin de mejorar la capacidad de generalización del modelo y reducir posibles sesgos. Sería valioso también incluir géneros musicales adicionales, como salsa, merengue, entre otros géneros, lo que permitiría evaluar el desempeño del modelo en contextos multicategoría. Por otro lado, para enriquecer el proceso de selección de características, se propone incorporar Coeficientes Cepstrales en las Frecuencias Melódicas (MFCCs), ampliamente utilizados en tareas de reconocimiento de audio debido a su capacidad para capturar el timbre de los instrumentos. Por último, se considera pertinente experimentar con otros algoritmos de clasificación como Redes Neuronales Convolucionales (CNN), que permiten aprender representaciones directamente a partir de los espectrogramas, lo cual podría robustecer la identificación automática de patrones sonoros y mejorar el rendimiento general del sistema de clasificación.

Bibliografía

- Academia Lab (2025). El método de welch. <https://academia-lab.com/enciclopedia/el-metodo-de-welch/>. Enciclopedia. Revisado el 23 de julio del 2025.
- Ambardar, A. (1995). *Procesamiento de Señales Analógicas y Digitales*. Thomson Learning.
- Aponte Mantilla, M. E. (2011). La historia del vallenato: discursos hegemónicos y disidentes.
- Banitalebi-Dehkordi, M. and Banitalebi-Dehkordi, A. (2014). Music genre classification using spectral analysis and sparse representation of the signals. *Journal of Signal Processing Systems*, 74:273–280.
- Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology*, 50(1):73–98.
- Benetos, E., Dixon, S., Giannoulis, D., Kirchhoff, H., and Klapuri, A. (2018). Automatic music transcription: An overview. *IEEE Signal Processing Magazine*, 36(1):20–30.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction*. John Wiley & Sons, New York.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.
- Brillinger, D. R. (1974). *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, New York.
- Caraballo, G. (2024). Egidio, la nota al cuadrado que necesitaba el vallenato. *El País*. Consultado el 3 de marzo de 2025.
- Carvalho, A. P. and Silva, J. L. (2012). Economic cycle analysis using singular spectrum analysis. *Journal of Economic Analysis*, 45:123–145.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.
- Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Cruz, D. A., Cristancho, C. C., and Camargo, J. E. (2019). Automatic identification of traditional colombian music genres based on audio content analysis and machine learning techniques. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 24th*

- Iberoamerican Congress, CIARP 2019, Havana, Cuba, October 28-31, 2019, Proceedings 24*, pages 646–655. Springer.
- Davis, S. and Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4):357–366.
- Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7):1895–1923.
- Dixon, S. (2007). Evaluation of the audio beat tracking system beatroot. *Journal of New Music Research*, 36(1):39–50.
- Echeverry, J. D., Lemus, C. G., and Orozco, Á. Á. (2007). Análisis de la densidad espectral de potencia en registros mer. *Scientia et technica*, 1(35).
- Escobar, J. (2005). *Colombia: Tierra de esperanza*. Editorial Cultural, Bogotá, Colombia.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182.
- III, J. O. S. (2011). *Spectral Audio Signal Processing*. W3K Publishing. <https://ccrma.stanford.edu/~jos/sasp/>.
- Jensen, J., Arnfast, J., and Heijenoort, M. (2006). Temporal feature integration for music genre classification. *IEEE Signal Processing Letters*, 13(6):324–327.
- Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer.
- Lartillot, O. and Toivainen, P. (2007). A matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects (DAFx)*, pages 237–244.
- Lartillot, O. and Toivainen, P. (2008). Mirtoolbox 1.0: Audio analysis tools for music information retrieval. *International Society for Music Information Retrieval Conference*, pages 23–27.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.
- Londoño, A. (1985). El currulao. *Educación Física y Deporte*, 7(1-2):78–90.
- Louppe, G. (2014). Understanding random forests: From theory to practice. Technical Report arXiv:1407.7502, arXiv.
- Luque, H. F. (2012). *La guitarra vallenata: adaptaciones de paseo y merengue vallenato para formatos de cuarteto de guitarra y guitarra solista*. PhD thesis, Universidad Pedagógica Nacional, Bogotá, Colombia. Recuperado del repositorio institucional.
- Martín, A., Martín, S., González, C., and Ay, E. (2001). Aplicación del análisis espectral al estudio de las masas forestales. In *Congresos Forestales*.

Bibliografía

- Martínez, C. A. (2018). Producción musical contemporánea utilizando las tradiciones del vallenato y el currulao.
- Mata Hernández, G., Sánchez Esquivel, V. M., and Gómez González, J. M. (2017). *Análisis de sistemas y señales*. Universidad Nacional Autónoma de México.
- Mendes, R. S., Ribeiro, H. V., Freire, F. C. M., Tateishi, A. A., and Lenzi, E. K. (2010). Universal patterns in sound amplitudes of songs and music genres. *arXiv preprint arXiv:1012.0142*.
- Peeters, G. (2004). A large set of audio features for sound description (similarity and classification) in the cuidado project. Technical report, IRCAM - Centre Pompidou.
- Rehan Akbani, Stephen Kwek, N. J. (2004). Applying support vector machines to imbalanced datasets. *European Conference on Machine Learning*, pages 39–50.
- Rey Bernal, L. D. (2019). Análisis de la polirritmia del currulao del pacífico colombiano. Acceso abierto.
- Rumsey, F. (2003). *Desktop Audio Technology: Digital Audio and MIDI Principles*. Focal Press.
- Scaringella, N., Zoia, G., and Mlynek, D. (2006). Automatic genre classification of music content. *IEEE Signal Processing Magazine*, 23(2):133–141.
- Scheirer, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, 103(1):588–601.
- Shumway, R. H. and Stoffer, D. S. (2017). Arima models. In *Time series analysis and its applications: with R examples*, pages 75–163. Springer.
- Sueur, J. et al. (2018). *Sound analysis and synthesis with R*. Springer.
- Toledo, J. A., Muñoz, A., and Castellanos, G. (2019). Music genre classification based on spectral features and machine learning techniques. *Journal of Audio Engineering Society*, 67(4):235–245.
- Trevor Hastie, Robert Tibshirani, J. F. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2 edition.
- Tzanetakis, G. and Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302.
- UNESCO (2015). El vallenato, música tradicional de la región del magdalena grande.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer.
- Vapnik, V. (1998). *Statistical Learning Theory*. Wiley, New York.
- Welch, P. D. (1967). *The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms*. IEEE.

A. Anexo