



INSTITUTO POLITÉCNICO NACIONAL  
ESCUELA SUPERIOR DE CÓMPUTO

**ESCOM**

*Trabajo Terminal:*

**“Clasificador Unimodal de  
Archivos Musicales”**

*2012 - B021*

*Que para cumplir con la opción de titulación curricular en la carrera de:*  
**“Ingeniería en Sistemas Computacionales”**

*Presentan*

Hernández Peralta Miguel Iván  
Rodríguez Covarrubias Isaac

Directores

M. en C. Mario Augusto Ramirez Morales

M. en C. Jaqueline Arzate Gordillo



ESCOM

México D.F. a 19 de Noviembre de 2013



INSTITUTO POLITÉCNICO NACIONAL  
ESCUELA SUPERIOR DE CÓMPUTO



No. de registro: TT 2012 B-021

Serie: Amarilla

Noviembre de 2013

Documento Técnico

**Clasificador Unimodal de Archivos Musicales**

Autores:

Hernández Peralta Miguel Iván<sup>[1]</sup>

Rodríguez Covarrubias Isaac<sup>[2]</sup>

**Resumen:**

En este reporte se presenta la documentación técnica del Trabajo Terminal 2012-B021 titulado **Clasificador Unimodal de Archivos Musicales**, cuyo objetivo es la clasificación de archivos de audio en alguna taxonomía mediante el análisis y procesamiento de las señales digitales de los archivos de audio, ahora considerando que mientras más avanza el tiempo, el repertorio musical crece de una manera extremadamente rápida y su organización manual suele ser muy tardada. Es por esto que apoyándonos de el procesamiento de los equipos de cómputo se pretende organizarlas mediante la taxonomía definida por su género musical.

**Palabras Clave:** Clasificación, Música, Géneros Musicales

Directores

M. en C Mario Augusto Ramirez Morales

M. en C. Jaqueline Arzate Gordillo

1 mikeivan21@hotmail.com

2 isaacirc2cm8@gmail.com



ESCUELA SUPERIOR DE CÓMPUTO  
SUBDIRECCIÓN ACADÉMICA  
DEPARTAMENTO DE FORMACIÓN INTEGRAL E INSTITUCIONAL  
COMISIÓN ACADÉMICA DE TRABAJOS TERMINALES



México, D.F. a 19 de Noviembre de 2013.

DR. FLAVIO ARTURO SÁNCHEZ GARFIAS  
PRESIDENTE DE LA COMISIÓN ACADÉMICA  
DE TRABAJOS TERMINALES  
P R E S E N T E

Por medio del presente, informamos que los alumnos que integran el TRABAJO TERMINAL 2012 B-021 titulado “Clasificador Unimodal de Archivos Musicales” concluyeron satisfactoriamente su trabajo.

El empastado del Reporte Técnico Final y el Disco Compacto (CD) fueron revisados y corregidos ampliamente por sus servidores: Mario Augusto Ramírez Morales y Jaqueline Arzate Gordillo, cubriendo el alcance y el objetivo planteados en el protocolo original y de acuerdo a los requisitos establecidos por la Comisión que Usted preside.

ATENTAMENTE

M. en C. Mario Augusto  
Ramírez Morales

M. en C. Jaqueline  
Arzate Gordillo

DIRECTORES DEL TRABAJO TERMINAL

## Advertencia

“Este documento contiene información desarrollada por la Escuela Superior de Cómputo del Instituto Politécnico Nacional, a partir de datos y documentos con derecho de propiedad y por lo tanto, su uso quedará restringido a las aplicaciones que explícitamente se conven- gan.”

La aplicación no convenida exime a la escuela su responsabilidad técnica y da lugar a las consecuencias legales que para tal efecto se determinen.

Información adicional sobre este reporte técnico podrá obtenerse en: La Subdirección Académica de la Escuela Superior de Cómputo del Instituto Politécnico Nacional, situada en Av. Juan de Dios Bátiz s/n Teléfono: 57296000 Extensión 52000 © 2013 - *Hernández Peralta Miguel Iván, Rodríguez Covarrubias Isaac*  
Todos los Derechos Reservados.

*"A mi hermano Eric por su amistad y cariño  
que me impulsan a ser mejor cada día para darle un buen ejemplo a seguir.  
A mis abuelitas por sus cuidados y cariños.  
A mis padres por su apoyo."*

Isaac Rodríguez

*"A mis padres. Por sus regaños y enseñanzas, por la fuerza  
que me han dado para no derrumbarme ante las adversidades de esta vida  
porque cuando he estado a punto de caer lo único que recibí  
de ustedes son sus consejos y su ayuda.  
Siempre los llevare conmigo. Gracias."*

Miguel Iván Hernández

# Agradecimientos

*Al Instituto Politécnico Nacional por abrirnos sus puertas para poder formar parte de él.  
Gracias*

*A nuestra escuela ESCOM “Escuela Superior de Computo”  
de la cual nos sentimos orgullosos de pertenecer y llevar acabo nuestros estudios en ella.  
Gracias.*

*”A todos los docentes que nos han formado a lo largo de nuestra vida escolar”  
Gracias.*

*A nuestros directores Mario Augusto Ramirez Morales y Jaqueline Arzate Gordillo, por su  
tiempo, entusiasmo, experiencia y conocimientos compartidos a lo largo de el desarrollo  
de este trabajo  
Gracias.*

Miguel Iván Hernández, Isaac Rodríguez

# Índice General

1	Introducción	1
1.1	Antecedentes	1
1.2	Planteamiento del Problema	2
1.3	Objetivo General	3
1.4	Objetivos Particulares	3
1.5	Justificación	3
2	Marco Teórico	5
2.1	El sonido	5
2.2	Los géneros musicales	5
2.2.1	Origen de los géneros musicales	8
2.3	El audio digital	18
2.4	Digitalización del sonido	19
2.4.1	Conceptos	19
2.4.2	Digitalización	19
2.5	Características del sonido	20
2.5.1	En el dominio del tiempo	20
2.5.2	En el dominio de la frecuencia	22
2.5.3	El Cepstrum	23
2.6	Archivos de sonido	25
2.6.1	Archivos de sonido con pérdida	25
2.6.2	Archivos de sonido sin pérdida	26

2.7	Archivo Wav . . . . .	27
2.7.1	Limitantes del formato Wav . . . . .	28
2.8	Redes Neuronales . . . . .	29
2.8.1	Aprendizaje No Supervisado . . . . .	30
2.8.2	Aprendizaje Supervisado . . . . .	30
2.9	Paradigma clasificadorio K-NN (K - Nearest Neighbour) . . . . .	31
2.9.1	Algoritmo K-NN básico [1] . . . . .	31
<b>3</b>	<b>Problemática</b>	<b>34</b>
3.1	Descripción de la problemática . . . . .	34
<b>4</b>	<b>Análisis</b>	<b>37</b>
4.1	Estudio de factibilidad . . . . .	37
4.1.1	Factibilidad Operacioinal . . . . .	37
4.1.2	Factibilidad Técnica . . . . .	39
4.1.3	Factibilidad Económica . . . . .	39
4.1.4	Factibilidad Legal . . . . .	40
4.2	Estudio de software y hardware . . . . .	40
4.2.1	Hardware . . . . .	40
4.2.2	Software . . . . .	41
4.2.3	Archivos de audio . . . . .	42
4.3	Análisis de requerimientos . . . . .	42
4.3.1	Requerimientos funcionales . . . . .	42
4.3.2	Requerimientos no funcionales . . . . .	42
4.4	Análisis de riesgos . . . . .	43
4.4.1	Objetivo del análisis de riesgos . . . . .	43
4.4.2	Identificación de riesgos . . . . .	43
4.4.3	Plan de riesgo . . . . .	44
<b>5</b>	<b>Productos o resultados esperados</b>	<b>45</b>
5.1	Metodología . . . . .	45
5.2	Definición de prototipos . . . . .	46

5.3	Diagrama de Bloques . . . . .	47
5.4	Arquitectura general de proyecto . . . . .	48
5.5	Diagrama de interacción de módulos . . . . .	49
<b>6</b>	<b>Prototipo I: Lectura de archivos WAV en un directorio</b>	<b>50</b>
6.1	1° Versión . . . . .	51
6.1.1	Análisis y diseño . . . . .	51
6.1.2	Funcionamiento . . . . .	57
<b>7</b>	<b>Prototipo II: Procesamiento de la señal digital de audio</b>	<b>59</b>
7.1	Requerimientos funcionales genéricos . . . . .	60
7.2	1° Versión . . . . .	61
7.2.1	Requerimientos no funcionales . . . . .	62
7.2.2	Diagrama de Casos de Uso . . . . .	63
7.2.3	Diagrama de Clases . . . . .	64
7.2.4	Diagrama de Actividades . . . . .	65
7.3	2° Versión . . . . .	67
7.3.1	Requerimientos no funcionales . . . . .	67
7.3.2	Diagrama de Casos de Uso . . . . .	68
7.3.3	Diagrama de Clases . . . . .	69
7.3.4	Diagrama de Actividades . . . . .	70
7.4	3° Versión . . . . .	72
7.4.1	Requerimientos no funcionales . . . . .	72
7.4.2	Diagrama de Casos de Uso . . . . .	73
7.4.3	Diagrama de Clases . . . . .	74
7.4.4	Diagrama de Actividades . . . . .	75
7.5	4° Versión . . . . .	76
7.5.1	Requerimientos no funcionales . . . . .	77
7.5.2	Diagrama de Casos de Uso . . . . .	78
7.5.3	Diagrama de Clases . . . . .	79
7.5.4	Diagrama de Actividades . . . . .	80

7.6	5° Versión . . . . .	81
7.6.1	Requerimientos no funcionales . . . . .	82
7.6.2	Diagrama de Casos de Uso . . . . .	84
7.6.3	Diagrama de Clases . . . . .	84
7.6.4	Diagrama de Actividades . . . . .	85
7.7	6° Versión . . . . .	87
7.7.1	Requerimientos no funcionales . . . . .	87
7.7.2	Diagrama de Casos de Uso . . . . .	89
7.7.3	Diagrama de Clases . . . . .	89
7.7.4	Diagrama de Actividades . . . . .	90
7.8	7° Versión . . . . .	92
7.8.1	Diagrama de Casos de Uso . . . . .	94
7.8.2	Diagrama de Clases . . . . .	94
7.8.3	Diagrama de Actividades . . . . .	95
7.9	8° Versión . . . . .	97
7.9.1	Requerimientos no funcionales . . . . .	97
7.9.2	Diagrama de Casos de Uso . . . . .	98
7.9.3	Diagrama de Clases . . . . .	99
7.9.4	Diagrama de Actividades . . . . .	100
7.10	9° Versión . . . . .	102
7.10.1	Requerimientos no funcionales . . . . .	102
7.10.2	Diagrama de Casos de Uso . . . . .	104
7.10.3	Diagrama de Clases . . . . .	104
7.10.4	Diagrama de Actividades . . . . .	105
<b>8</b>	<b>Prototipo III: Clasificador de archivos musicales</b>	<b>107</b>
8.1	1° Versión: Vecinos más cercanos. . . . .	107
8.2	2° Versión: Máquina de soporte vectorial. . . . .	109
<b>9</b>	<b>Prototipo IV: Prototipo Integrador</b>	<b>111</b>
9.1	Requerimientos funcionales genéricos . . . . .	111

9.2	1° Versión. . . . .	112
9.2.1	Requerimientos no funcionales . . . . .	112
9.2.2	Diseño de la interfaz . . . . .	113
<b>10</b>	<b>Pruebas de clasificación</b>	<b>118</b>
10.1	1° Prueba . . . . .	118
10.1.1	Contexto . . . . .	118
10.1.2	Resultados . . . . .	119
10.2	2° Prueba . . . . .	122
10.2.1	Contexto . . . . .	122
10.2.2	Resultados . . . . .	122
10.2.3	Conclusiones . . . . .	123
10.3	3° Prueba . . . . .	123
10.3.1	Contexto . . . . .	123
10.3.2	Resultados . . . . .	124
10.4	4° Prueba . . . . .	124
10.4.1	Contexto . . . . .	124
10.4.2	Resultados . . . . .	125
10.5	5° Prueba . . . . .	126
10.5.1	Contexto . . . . .	126
10.5.2	Resultados . . . . .	126
10.6	6° Prueba . . . . .	127
10.6.1	Contexto . . . . .	127
10.6.2	Resultados . . . . .	127
10.7	7° Prueba . . . . .	128
10.7.1	Contexto . . . . .	128
10.7.2	Resultados . . . . .	128
10.8	8° Prueba . . . . .	129
10.8.1	Contexto . . . . .	129
10.8.2	Resultados . . . . .	130
10.9	9° Prueba . . . . .	143

10.9.1 Contexto . . . . .	143
10.9.2 Resultados . . . . .	144
10.1010° Prueba . . . . .	144
10.10.1 Contexto . . . . .	144
10.10.2 Resultados . . . . .	145
10.1111° Prueba . . . . .	145
10.11.1 Contexto . . . . .	145
10.11.2 Resultados . . . . .	148
10.1212° Prueba . . . . .	148
10.12.1 Contexto . . . . .	148
10.12.2 Resultados . . . . .	148
10.1313° Prueba . . . . .	149
10.13.1 Contexto . . . . .	149
10.13.2 Resultados . . . . .	149
10.1414° Prueba . . . . .	150
10.14.1 Contexto . . . . .	150
10.14.2 Resultados . . . . .	150
10.1515° Prueba . . . . .	151
10.15.1 Contexto . . . . .	151
10.15.2 Resultados . . . . .	151
10.1616° Prueba . . . . .	152
10.16.1 Contexto . . . . .	152
10.16.2 Resultados . . . . .	152
10.1717° Prueba . . . . .	153
10.17.1 Contexto . . . . .	153
10.17.2 Resultados . . . . .	153
10.1818° Prueba . . . . .	154
10.18.1 Contexto . . . . .	154
10.18.2 Resultados . . . . .	156
10.1919° Prueba . . . . .	157
10.19.1 Contexto . . . . .	157

10.19.2 Resultados . . . . .	158
<b>11 Referencias musicales</b>	<b>159</b>
11.1 Pistas de Entrenamiento . . . . .	159
11.1.1 Género clásica. . . . .	159
11.1.2 Género electrónica. . . . .	161
11.1.3 Género hip-hop. . . . .	163
11.1.4 Género jazz. . . . .	164
11.1.5 Género pop. . . . .	165
11.1.6 Género rock. . . . .	166
11.1.7 Género salsa. . . . .	167
11.2 Pistas de Pruebas . . . . .	169
11.2.1 Género clásica. . . . .	169
11.2.2 Género Electrónica. . . . .	169
11.2.3 Género Hip-Hop. . . . .	169
11.2.4 Género Jazz. . . . .	170
11.2.5 Género Pop. . . . .	170
11.2.6 Género Rock. . . . .	171
11.2.7 Género Salsa. . . . .	171
11.3 Referencias de entrenamiento masivo . . . . .	172
<b>Referencias</b>	<b>173</b>

# Índice de Imágenes

2.5.1 Cepstrum de un fragmento de señal. [8]	24
2.8.1 Diagrama de bloque de una red neuronal	31
2.9.1 Pseudocódigo para el clasificador K-NN	33
2.9.2 Ejemplo de aplicación del algoritmo K-NN básico	33
5.1.1 Proceso de la metodología del Prototipo Evolutivo.	46
5.3.1 Diagrama de Bloques	48
5.4.1 Diagrama de arquitectura general.	48
5.5.1 Diagrama de interacción de módulos	49
6.1.1 Diagrama a bloques Prototipo 1.	51
6.1.2 Diagrama de clases.	52
6.1.3 Diagrama de casos de uso.	53
6.1.4 Diagrama de actividades de ejecución y copia de archivos.	54
6.1.5 Diagrama de actividades de lectura y obtención de rutas absolutas de archivos.	55
6.1.6 Diagrama de actividades de la obtención de la señal digital.	56
6.1.7 Diagrama de actividades de la obtención de la señal digital.	57
6.1.8 Funcionamiento Porototipo 1	58
9 Diagrama a bloques del Prototipo II.	60
7.1.1 Diagrama de secuencia genérica del Prototipo II.	61
7.2.1 Diagrama de casos de uso del Prototipo II versión 1.	64
7.2.2 Diagrama de clases del Prototipo II versión 1.	65

7.2.3 Diagrama de actividades de la extracción de características del Prototipo II versión 1. . . . .	66
7.3.1 Diagrama de casos de uso del Prototipo II versión 2. . . . .	69
7.3.2 Diagrama de clases del Prototipo II versión 2. . . . .	70
7.3.3 Diagrama de actividades de la extracción de características del Prototipo II versión 2. . . . .	71
7.4.1 Diagrama de casos de uso del Prototipo II versión 3. . . . .	74
7.4.2 Diagrama de clases del Prototipo II versión 3. . . . .	75
7.4.3 Diagrama de actividades de la extracción de características del Prototipo II versión 3. . . . .	76
7.5.1 Diagrama de casos de uso del Prototipo II versión 4. . . . .	79
7.5.2 Diagrama de clases del Prototipo II versión 4. . . . .	80
7.5.3 Diagrama de actividades de la extracción de características del Prototipo II versión 4. . . . .	81
7.6.1 Diagrama de casos de uso del Prototipo II versión 5. . . . .	84
7.6.2 Diagrama de clases del Prototipo II versión 5. . . . .	85
7.6.3 Diagrama de actividades de la extracción de características del Prototipo II versión 5. . . . .	86
7.7.1 Diagrama de casos de uso del Prototipo II versión 6. . . . .	89
7.7.2 Diagrama de clases del Prototipo II versión 6. . . . .	90
7.7.3 Diagrama de actividades de la extracción de características del Prototipo II versión 6. . . . .	91
7.8.1 Diagrama de casos de uso del Prototipo II versión 7. . . . .	94
7.8.2 Diagrama de clases del Prototipo II versión 7. . . . .	95
7.8.3 Diagrama de actividades de la extracción de características del Prototipo II versión 7. . . . .	96
7.9.1 Diagrama de casos de uso del Prototipo II versión 8. . . . .	99
7.9.2 Diagrama de clases del Prototipo II versión 8. . . . .	100
7.9.3 Diagrama de actividades de la extracción de características del Prototipo II versión 8. . . . .	101
7.10.1 Diagrama de casos de uso del Prototipo II versión 9. . . . .	104

7.10.2 Diagrama de clases del Prototipo II versión 9. . . . .	105
7.10.3 Diagrama de actividades de la extracción de características del Prototipo II versión 9. . . . .	106
8.1.1 Diagrama de clases del Prototipo III KNN. . . . .	109
8.2.1 Diagrama de clases del Prototipo III SVM. . . . .	110
9.2.1 Diseño de la pantalla principal del Prototipo IV versión 1. . . . .	114
9.2.2 Mensaje para seleccionar el género a entrenar. . . . .	115
9.2.3 Mensaje que muestra que se ha clasificado. . . . .	116
9.2.4 Mensaje que muestra que se han transferido los archivos. . . . .	117
10.1.1 Características en el tiempo del género clásica. . . . .	119
10.1.2 Características en el tiempo del género electrónica. . . . .	119
10.1.3 Características en el tiempo del género hip-hop. . . . .	120
10.1.4 Características en el tiempo del género jazz. . . . .	120
10.1.5 Características en el tiempo del género pop. . . . .	121
10.1.6 Características en el tiempo del género rock. . . . .	121
10.1.7 Características en el tiempo del género salsa. . . . .	122
10.2.1 Matriz de resultados prueba 2. . . . .	123
10.3.1 Matriz de resultados prueba 3. . . . .	124
10.4.1 Matriz de resultados prueba 4. . . . .	125
10.5.1 Tabla que muestra los resultados de la prueba con SVM prueba 5. . . . .	126
10.5.2 Tabla que muestra los resultados de la prueba con KNN prueba 5. . . . .	127
10.6.1 Matriz de resultados prueba 6. . . . .	128
10.7.1 Matriz de resultados prueba 7. . . . .	129
10.8.1 Gráfica del género clásica. . . . .	130
10.8.2 Gráfica del género electrónica. . . . .	131
10.8.3 Gráfica del género hip-hop. . . . .	131
10.8.4 Gráfica del género jazz. . . . .	132
10.8.5 Gráfica del género pop. . . . .	132
10.8.6 Gráfica del género rock. . . . .	133

10.8.7	Gráfica del género salsa. . . . .	133
10.8.8	Gráfica del promedio de muestras del género clásica. . . . .	134
10.8.9	Gráfica del promedio de muestras del género electrónica. . . . .	135
10.8.10	Gráfica del promedio de muestras del género hip-hop. . . . .	135
10.8.11	Gráfica del promedio de muestras del género jazz. . . . .	136
10.8.12	Gráfica del promedio de muestras del género pop. . . . .	136
10.8.13	Gráfica del promedio de muestras del género rock. . . . .	137
10.8.14	Gráfica del promedio de muestras del género salsa. . . . .	137
10.8.15	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género clásica. . . . .	138
10.8.16	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género electrónica. . . . .	139
10.8.17	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género hip-hop. . . . .	139
10.8.18	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género jazz. . . . .	140
10.8.19	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género pop. . . . .	140
10.8.20	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género rock. . . . .	141
10.8.21	Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género salsa. . . . .	141
10.8.22	Gráfica comparativa del promedio de muestras reducido de 1024 a 16 elementos de todos los géneros. . . . .	142
10.8.23	Gráfica comparativa del género rock y hiphop. . . . .	143
10.9.1	Gráfica de resultados contemplando 1024 características prueba 9. . . . .	144
10.9.2	Gráfica de resultados contemplando 2048 características prueba 9. . . . .	144
10.10.1	Gráfica de resultados contemplando 8 características prueba 10. . . . .	145
10.10.2	Gráfica de resultados contemplando 16 características prueba 10. . . . .	145
10.11.1	Primer árbol de decisión. . . . .	146
10.11.2	Segundo árbol de decisión. . . . .	147

10.11.3 Tercer árbol de decisión. . . . .	147
10.11.4 Gráfica de resultados realizando un árbol de desición prueba 11. . . . .	148
10.12.1 Gráfica de resultados contemplando 16 características. . . . .	149
10.13.1 Gráfica de resultados contemplando 1024 características y 500 muestras de entrenamiento por género. . . . .	150
10.14.1 Gráfica de resultados contemplando 512 características. . . . .	151
10.15.1 Gráfica de resultados contemplando el módulo de la FFT. . . . .	152
10.16.1 Gráfica de resultados contemplando la Fase de la FFT. . . . .	153
10.17.1 Gráfica de resultados contemplando la fase y el módulo de la FFT. . . . .	154
10.18.1 Gráfica de FFT de 1024 puntos promedio por género. . . . .	155
10.18.2 Primeros puntos con mayor diferencia de la FFT promedio de 1024 ele- mentos. . . . .	156
10.18.3 Gráfica de resultados contemplando los primeros 50 elementos de la FFT. . . . .	157
10.19.1 Gráfica de resultados los tres mejores géneros. . . . .	158

# Índice de tablas

2.2.1 Tabla con características de los géneros musicales. . . . .	7
2.9.1 Notación para el paradigma K-NN. . . . .	32
3.1.1 Tabla de sistemas similares y solución propuesta . . . . .	36
4.1.1 Tabla de costo de horas de trabajo. . . . .	40
4.2.1 Tabla de hardware requerido. . . . .	41
4.2.2 Tabla de software requerido. . . . .	41
4.4.1 Tabla de riesgos. . . . .	43
4.4.2 Tabla con plan de riesgos. . . . .	44
5.2.1 Prototipos que integrarán el sistema. . . . .	47
6.1.1 Tabla de requerimientos funcionales del Prototipo 1. . . . .	51
6.1.2 Tabla de requerimientos no funcionales del Prototipo 1. . . . .	52
7.1.1 Tabla de requerimientos funcionales genéricos del Prototipo II. . . . .	61
7.2.1 Tabla de requerimientos no funcionales de la versión 1 del Prototipo II. . . . .	63
7.3.1 Tabla de requerimientos no funcionales de la segunda versión del Prototipo II. . . . .	68
7.4.1 Tabla de requerimientos no funcionales del Prototipo II versión 3. . . . .	73
7.5.1 Tabla de requerimientos no funcionales del Prototipo II versión 4. . . . .	78
7.6.1 Tabla de requerimientos no funcionales del Prototipo II versión 5. . . . .	83
7.7.1 Tabla de requerimientos no funcionales del Prototipo II versión 6. . . . .	88

7.8.1	Tabla de requerimientos no funcionales del Prototipo II versión 7. . . . .	93
7.9.1	Tabla de requerimientos no funcionales del Prototipo II versión 8. . . . .	98
7.10.	Tabla de requerimientos no funcionales del Prototipo II versión 9. . . . .	103
9.1.1	Tabla de requerimientos funcionales genéricos del Prototipo IV. . . . .	112
9.2.1	Tabla de requerimientos no funcionales de la versión 1 del Prototipo 2. . . .	113

*”Hay hombres que luchan un día y son buenos. Hay otros que luchan un año y son mejores. Hay quienes luchan muchos años y son muy buenos. Pero hay los que luchan toda la vida: esos son los imprescindibles”*

Bertolt Brecht

# 1

## Introducción

### 1.1 Antecedentes

Desde la antigüedad conforme el ser humano fue evolucionando, y volviéndose sedentario tuvo la necesidad de poseer cosas útiles y conservarlas. Conforme fue avanzando el tiempo comenzó a generar procesos repetitivos, y de esta manera crear su aprendizaje. Se dió cuenta de que podía realizar sonidos y mediante ellos comunicarse con otros seres humanos y creo el lenguaje, dándole significados universales a los sonidos, también observo que podía realizar sonidos con diferentes herramientas y fue así como desarrolló música.

La música nace con el ser humano y juega un papel muy importante en su entorno cultural y emocional, debido a que es una manera de expresar sentimientos. Por tanto la música es también considerada un arte por ser una manifestación del humano que lo ha acompañado a lo largo de su existencia.

Durante siglos, la música llegaba al público directamente de manos de aquellos que la interpretaban. Con la aparición, a finales del siglo XIX, de los primeros aparatos de grabación, y la evolución de las técnicas de edición y grabación musicales, en el siglo XX, la comercialización de música grabada se convirtió en un nuevo y potente medio de difusión musical es por esto que en la actualidad se cuenta con un repertorios musicales muy amplios.

Estos desarrollos teconológicos cuentan con herramientas que nos permiten grabar, almacenar y reproducir sonidos y música en instantes, permiten a los usuarios poseer la música de su agrado en grandes cantidades en archivos musicales digitales, Pero para el ser humano no es suficiente almacenar lo que le agrada o le es de utilidad, ya que trata de mantener un ordenar, es por eso que la clasificación de objetos o elementos ha sido de vital importancia para su estudio, anteriormente la realizába mediante la percepción de sus sentidos, creando estándares o patrones abstractos para segmentar un número de elementos, pero actualmente el análisis de un equipo de cómputo es muy superior a la percepción humana.

## 1.2 Planteamiento del Problema

Años atrás la música se comenzó a producir y distribuir en formato digital, y es hoy en día una forma muy común de almacenarla, pero debido a la gran diversidad de géneros musicales existentes, resulta difícil mantenerla clasificada constantemente, ya que las colecciones cambian y crecen día a día.

En 2004 el repertorio de música digital era de apenas 1 millón de unidades hoy en día existe un repertorio de 18 millones de archivos musicales disponibles en la red en diversas aplicaciones [11]

Esto nos indica que el repertorio de música digital continuará en constante crecimiento por tanto, el problema radica en mantener organizada la biblioteca digital de música de los usuarios.

### 1.3 Objetivo General

Clasificación de archivos de audio en alguna taxonomía mediante el análisis de sus características cuantitativas.

### 1.4 Objetivos Particulares

- ⇒ Analizar las características cuantitativas de la señal de audio de los archivos musicales para poder determinar las similitudes existentes entre géneros musicales y posteriormente mediante estas similitudes lograr su clasificación.
- ⇒ Buscar características cuantitativas de la señal digital de audio que nos permitan realizar una clasificación.
- ⇒ Entrenar la aplicación con archivos musicales de los cuales el usuario tiene identificado el género musical al que pertenecen.
- ⇒ Lograr reconocer el género musical de los archivos de audio previo a la etapa de entrenamiento realizada.

### 1.5 Justificación

En la actualidad se produce y se distribuye una gran cantidad de música en formatos digitales, y se almacena en grandes cantidades, pero al conservarla surge la necesidad de clasificarla para su mejor organización.

¿Por qué este trabajo tendrá la complejidad suficiente para obtener el título de Ingeniero en Sistemas Computacionales? Este trabajo requiere conocimientos previamente adquiridos

durante nuestra estancia en la Escuela Superior de Cómputo, tales como:

- ⇒ Análisis de Señales
- ⇒ Matemáticas Avanzadas para la Ingeniería
- ⇒ Ingeniería de Software
- ⇒ Análisis de Algoritmos
- ⇒ Inteligencia Artificial

# 2

## Marco Teórico

### 2.1 El sonido

El sonido, en física, es cualquier fenómeno que involucre una propagación en forma de ondas elásticas (sean audibles o no), generalmente a través de un fluido (u otro medio elástico) que esté generando el movimiento vibratorio de un cuerpo. [19]

### 2.2 Los géneros musicales

La música es antes que todo un arte, que combina los sonidos y el silencio conforme a los principios de la melodía, la armonía y el ritmo, siendo estos elementos esenciales. Al mismo tiempo es una ciencia, por lo cual debe ser apreciada emocionalmente y comprendida intelectualmente. [7]

El concepto de género musical se emplea en música para clasificar las obras musicales, esta clasificación se puede hacer de distinta forma dependiendo de los criterios que se utilicen para realizarla, una obra musical puede pertenecer por lo tanto a varios géneros al mismo tiempo. [10]

Algunos de los criterios usados para la clasificación por género pueden ser específicamente musicales como el ritmo, la instrumentación, las características armónicas o melódicas o su estructura, y también basarse en características no musicales como la región geográfica de origen, el periodo histórico, el contexto sociocultural u otros aspectos más amplios de una determinada cultura. [20]

En la tabla 2.2.1 se describen brevemente algunas de las características de los géneros musicales que se pretende el Clasificador Unimodal de Archivos maneje.

Género	Origen	Ritmo	Instrumentos
Rock	Surgido en 1950 y creó una nueva cultura popular	Es de construcción tradicional sobre ritmos simples en una métrica 4/4, con una batería repetitiva en los tiempos dos y cuatro. [18]	Guitarra eléctrica rítmica, eléctrica solista, bajo eléctrico y batería.
Pop	Comienza durante los años 50. [6]	Ritmo constante y notable con una especial atención a las melodías pegadizas y un coro que contrasta melódica, rítmica y armónicamente con el verso. [14]	Guitarra, bajo eléctrico, batería y teclados.
Jazz	Nacido a finales del siglo XIX	Caracterizado por la improvisación con un sonido y frasea que reflejan la personalidad de los músicos ejecutantes, con un ritmo especial conocido como swing. [2]	Piano, trompeta, trombón, clarinete, saxofón, contrabajo, bajo eléctrico, batería, teclado, batería.
Electrónica	Surgida en 1952, y a inicios de los ochentas surge el house y el tecno.	El ritmo está dispuesto en beats (equivalente rítmico del bit digital) o golpes, con diferencias entre sus subgéneros por la medida de beats por minuto (bpm). [4]	Instrumentos electrónicos y tecnología musical electrónica como el theremin, el sintetizador de sonido y los equipos de cómputo. [22]
Tropical	La música tropical entro en Chile por primera vez en la década de 1930 con la guaracha.	Música vocal e instrumental asociada con la música caribeña. (salsa, merengue, cumbia y bachata). Ritmos sincopados y se basa en las formas musicales tradicionales.	Su sonido se basa en instrumentos como timbales, congas, bongo, los tambores, guache, maracas.
Clásica	El término música clásica aparece por primera vez en 1836 en el Oxford English Dictionary. citeForm14		Violín, cello, violas, contraBajos, clarinete, Oboe, fagot, trompeta, trombom, tuba, corno frances, tambor, flauta, timbales, platillo.
Hip hop	La música hip-hop o rap surgió en Estados Unidos a finales de los años 1960. [21]	La mayoría de los ritmos de hip-hop tienen bambos sincopados y una caja o palmada; muchos ritmos suelen acompañarse por melodías sencillas y repetitivas de sintetizador.	Pedazos de música vieja de los 60s y 70s, Soul, Funk, Jazz.

**Tabla 2.2.1:** Tabla con características de los géneros musicales.

### 2.2.1 Origen de los géneros musicales

Los generos musicales occidentales tienen sus orígenes en Grecia, donde desempeñó un papel moralizador y pedagógico desde la época en que se escribieron los textos homéricos (s. X a.C.).

musica

En el primer milenio de la era cristiana, la música fue cultivada exclusivamente por la iglesia. Hacia finales del siglo VI se refundió el llamado antiguo romano, y a este canto, adoptado de forma oficial por la iglesia católica, se le llamó desde entonces canto gregoriano extendiéndose por los diversos países europeos.

En el siglo IX, ya aparecen referencias al *órganum*, que era una forma medieval de escritura polifónica basada en una melodía (canto llano) sobre la que se superponía otra melodía.

A finales del siglo XI florece la lírica galante (no religiosa) de los trovadores provenzales, que inmediatamente se extiende al norte, los troveros en Francia y los *minnesinger* en Alemania y luego a Italia, España y Portugal.

A principios del 1200 surge el motete (canto a varias voces con distintas letras).

En el siglo XIV floreció en Francia y se extendió a Italia el *ars nova*, con repertorio profano (rondós, baladas, y madrigales), apareciendo músicos como Phillippe de Vitry o Francesco Landino.

En el 1400 se desarrolla la polifonía en la escuela franco-flamenca, a la que pertenecen Guillaume Dufay, Guilles Binchois y su máximo representante, Josquin des Prés. El renacimiento se inició como una prolongación de la etapa anterior, así nacieron los repertorios instrumentales.

La música se enriqueció verdaderamente en el periodo barroco con la ópera, la cantata, y el oratorio. Tres fueron los centros musicales: Italia, Alemania y Francia. Figuras señeras de la música, como Vivaldi, Haendel y Bach, aparecen en este periodo.

El clasicismo, fue el periodo que le siguió, abarcando la segunda mitad del siglo XVIII, destacando las figuras alemanas de Gluck, Haydn y Mozart. El periodo clásico aportó el mejoramiento de la técnica en todos sus aspectos.

En el siglo XIX, el romanticismo musical significó una libertad compositiva, desarrollándose los grandes géneros sinfónicos, abundando genios como: Beethoven, Schubert, Schumann, Chopin y Liszt. En un romanticismo tardío (si lo tomamos con humor), los nacionalismos hicieron su entrada en la música, destacando como representantes Brahms, Dvorak, Tchaikovski, Grieg, Verdi y Wagner. La ruptura con el romanticismo y la incorporación de innovaciones, se produjo en el último cuarto de siglo, con compositores como Debussy, Mahler, entre otros.

En el siglo XX, Schönberg aporta el atonalismo, imponiéndose las vanguardias estéticas: impresionismo, expresionismo, neoclasicismo, etc. Compositores dignos de mencionar son: Stravinski, Bártok, Prokófiev, Manuel de Falla. Cabe señalar la importancia de las emisoras de radio sobre la evolución y difusión de los géneros musicales en el mundo.

La proliferación musical de tendencias, fue sin duda la tónica del siglo XX, así tuvieron cabida géneros musicales como la música serial (desde 1923, basada en series matemáticas de alturas), la música concreta (desde 1948, basada en el procesamiento de grabaciones en cinta), la música electrónica (desde 1956, basada en la generación de sonido por medio de osciladores), la música electroacústica (combinación de las dos anteriores), la música aleatoria (John Cage), el postserialismo (Kagel, Hidalgo) y la música por ordenador.

## Clásica

La música clásica corresponde únicamente a la música culta compuesta en el período clásico, de mediados del siglo XVIII hasta principios del XIX. El término “música clásica” se ha extendido también a la música seria compuesta entre el medioevo y el presente.

En general se le llama “clásica”, pero los que saben, dicen que tal nombre se debe reservar para tal música, pero sólo para la que se hizo en un tiempo determinado, que se conoce como el período clásico de la música, aquel periodo en el que se destacaron Haydn, Mozart y Beethoven.

Puede ser llamada también música “culta”; pero en el concepto moderno, toda música es producto de una cultura, entendiendo por ésta, el conjunto de ideas, creencias religiosas, ciencias, artes y costumbres que forman y caracterizan el estado social de un pueblo o una raza.

De todas formas, lo que define a este estilo musical es excluyente. Se trata de aquella que no es popular ni folklórica. Es realizada por pocos, con la idea de que le guste a muchos, pero estos no suelen ser tantos en general. Sus autores y ejecutores han estudiado una larga carrera en escuelas especiales, conservatorios, y sus oyentes, en general han sido inducidos a gustarla por tradición familiar, que luego puede cultivarse. Requiere de cierta iniciativa personal para llegar a ella. Algo que si es definitivo es que ha podido trascender las fronteras del tiempo y el espacio; es universal. El mensaje que transmite la música clásica ha evolucionado en forma paralela al pensamiento occidental. Es su característica mas notable: no es estática, esta en continuo recambio de formas y modos, buscando siempre nuevos lenguajes, nuevas formas de expresión, cosa que no ocurre en la música de otras culturas.

A pesar de los limitantes de no ser folklórica ni popular, la música clásica está confinada a un ámbito geográfico, el de la cultura occidental, es decir, los pueblos europeos y sus herederos culturales, primero los americanos y después algunos otros. Hay también una

limitante temporal, refiriéndonos con el término “música clásica” a aquella creada a partir del Renacimiento y que representa el fenómeno cultural más brillante que se ha dado en la historia de la humanidad, solo comparable, por sus alcances, a la ciencia de la Europa posrenacentista.

Sus orígenes radican en las músicas medievales y renacentistas, quedando estas excluidas de la clasificación “clásica”. También se excluyen la oriental y la africana, sin desconocer sus valores y aportes. La música clásica occidental nos dice de estados anímicos muy profundamente arraigados en los europeos y sus herederos espirituales.

### Electrónica

Música electrónica es aquel tipo de música que emplea para su producción e interpretación instrumentos electrónicos y tecnología musical electrónica. En general, puede distinguirse entre el sonido producido utilizando medios electromecánicos de aquel producido utilizando tecnología electrónica. Ejemplos de dispositivos que producen sonido electromecánicamente son el telarmonio, el órgano Hammond y la guitarra eléctrica. La producción de sonidos puramente electrónica puede lograrse mediante aparatos como el theremin, el sintetizador de sonido y el ordenador.

La música electrónica se asoció en su día exclusivamente con una forma de música culta occidental, pero desde finales de los años 1960, la disponibilidad de tecnología musical a precios asequibles permitió que la música producida por medios electrónicos se hiciera cada vez más popular. En la actualidad, la música electrónica presenta una gran variedad técnica y compositiva, abarcando desde formas de música culta experimental hasta formas populares como la música electrónica de baile.

### Hip-hop

El hip hop o hip-hop es un movimiento artístico que surgió en Estados Unidos a finales de los años 1960 en las comunidades afroamericanas y latinoamericanas de barrios populares neoyorquinos como Bronx, Queens y Brooklyn. El DJ Afrika Bambaataa subrayó los cuatro

pilares de la cultura hip hop: MC, DJ, breakdance y grafiti. Otros elementos incluyen el beatboxing.

La música hip hop o rap es un género musical que se desarrolló como parte de la cultura hip hop, y se define por cuatro elementos estilísticos: MC, DJ turntablism, sampling (o sintetización de sonido) y beatboxing. El hip hop comenzó en el South Bronx de Nueva York a comienzos de los años 1970. El término rap suele utilizarse como sinónimo de hip hop, si bien hip hop también hace referencia a las prácticas de toda esta subcultura.

### Jazz

El jazz posee una historia definida y una evolución estilística específica, es una forma musical nacida hacia 1900 en Estados Unidos. En su desarrollo ha tomado elementos de la música folclórica; a su vez la música popular los ha tomado del jazz.

Desde sus inicios, el jazz se ha ramificado en muchos subestilos, lo que hace difícil lograr encontrar una descripción única que se adapte a todos ellos con fiabilidad absoluta. Sin embargo, pueden hacerse algunas generalizaciones, teniendo en cuenta que en todos los casos hay excepciones.

Según el estilo que se ha elegido, los intérpretes de jazz improvisan dentro de las convenciones posibles. Por lo general la improvisación es acompañada de una progresión de acordes de una canción popular o una composición original que se repiten. Los instrumentistas imitan los estilos vocales negros, incluso el uso de glisandos y portamentos (sonidos arrastrados de una nota a otra), las ligeras variaciones de tono, y los efectos sonoros tales como gruñidos y gemidos.

La voluntad de crear un sonido personal de color tonal con un sentido del ritmo y forma individual, y con un estilo propio de ejecución ha llevado a los músicos a la utilización de ritmos que se caracterizan por una sincopación constante (los acentos aparecen en momen-

tos inesperados del compás ) y también por el swing, un pulso rítmico único del jazz que se manifiesta en las complejas relaciones entre notas largas y breves. Las partituras escritas, si existen, se usan tan sólo como guías de la estructura dentro de la cual se desarrolla la improvisación. La instrumentación típica comienza con una sección rítmica, formada por el piano, el contrabajo, la batería y una guitarra opcional, a la que se pueden añadir instrumentos de viento. En las big bands los vientos se agrupan en tres secciones: saxofones, trombones y trompetas.

Teniendo algunas excepciones dentro de sus subestilos, el jazz en general se basa en la adaptación de infinitas melodías a algunas progresiones de determinados acordes. El músico improvisa nuevas melodías que se ajustan a la progresión de los acordes, y éstos se repiten tantas veces como se desee a medida que se incorpora cada nuevo solista.

Es claro que para la improvisación de jazz se usan piezas cuyas pautas formales son muy distintas, hay dos estructuras, en particular, que se usan con más frecuencia en sus temas. Una es la forma AABA de los estribillos de las canciones populares, que consta de 32 compases divididos en cuatro secciones de ocho compases cada una: la sección A, la repetición de la sección A, la sección B ( el puente, que suele comenzar en una tonalidad nueva ), y la repetición de la sección A. La segunda forma tiene hondas raíces en la música folclórica de la comunidad negra estadounidense, el blues de 12 compases. A diferencia de la forma de 32 compases en AABA, los blues tienen una progresión de acordes casi uniforme.

Las raíces del jazz radican en el eclecticismo musical de los afroamericanos. En esta tradición sobreviven elementos de la música del África occidental, de las formas musicales de la comunidad negra del Nuevo Mundo, de la música popular y clásica europea de los siglos XVII y XIX y de formas musicales populares posteriores que han influido en la música negra o que son obras de compositores negros. Los estilos vocales se encuentran entre los rasgos africanos, y destacan por una gran libertad de coloración vocal, la tradición de la improvisación, las pautas de pregunta y respuesta, y la complejidad rítmica, tanto en la síncope de líneas melódicas individuales como en los ritmos complejos que tocan los distintos

miembros de un conjunto. Los cantos que acompañaban el trabajo, las canciones de las nanas, posteriormente los cánticos espirituales y el blues eran otras formas de música que fueron asimiladas a los distintos estilos de jazz.

La música europea ha aportado estilos y formas específicas: himnos, marchas, valeses, cuadrillas y otras músicas de baile, de teatro, y de ópera italiana, así como elementos teóricos en especial la armonía, un vocabulario de acordes y la relación con la forma musical.

Entre los elementos negros de la música popular que han contribuido al desarrollo del jazz se incluyen la música de banjo de los minstrel shows ( derivados de la música de banjo de los esclavos), los ritmos sincopados de influencia negra procedentes de la música latinoamericana (que se escuchaba en las ciudades sureñas de Estados Unidos ), los estilos de pianola de los músicos de las tabernas del Medio Oeste, y las marchas e himnos interpretados por las bandas de metales de negros a finales del siglo XIX. En estos años surgió otro género que ejerció una poderosa influencia; se trataba del ragtime, música que combinaba muchos elementos, incluidos los ritmos sincopados ( originarios de la música de banjo y otras fuentes negras ) y los contrastes armónicos y las pautas formales de las marchas europeas. A partir de 1910 el director de orquesta W. C. Handy tomó otra forma, el blues, y la llevó más allá de su tradición precedente estrictamente oral, con la publicación de sus originales blues. En las manos de los músicos de jazz, sus blues encontraron en la década de 1920 a quien sería quizá su mejor intérprete: la cantante Bessie Smith, que grabó muchos de ellos.

La fusión de estas múltiples influencias en el jazz resulta difícil de reconstruir, dado que esto ocurrió antes de que el fonógrafo pudiera ofrecer testimonios fiables.

## Pop

La música pop es un estilo que ha logrado dominar las listas de ventas desde su nacimiento, a principios de la década de 1960, y que ha recibido etiquetas de toda clase, dada su extrema facilidad para adoptar variadas influencias ajenas. Los temas pop se basan en la repetición

de melodías en torno a una construcción clásica de estrofa, estribillo, estrofa, estribillo, puente y estribillo.

Siendo sus orígenes en el siglo XIX, no fue hasta el siglo XX, con la creación de la industria discográfica, cuando se convirtió como tantos otros géneros, en una realidad de consumo para la audiencia mundial, logrando desarrollar todo su potencial.

Fue influenciado en sus comienzos por el rock y el doo-wop, pero la música pop se ha basado siempre en la composición y la producción, y no en la actitud o el sentimiento propios del rock 6 roll y del blues, dando lugar a la llamada “cultura single” (o sencillo); grabaciones de un único tema que se alternan en las listas de éxitos. Durante la década de 1960 estuvo representado por The Beatles, The Beach Boys, The Everly Brothers o el productor Phil Spector.

Tras absorber la psicodelia de finales de esa década, compositores como Burt Bacharach, The Carpenters, Gilbert O’Sullivan o el dúo Simon and Garfunkel otorgaron al pop cierta profundidad y madurez, acercándose así a un público más adulto.

A pesar de ello, en la década de 1980 los creadores de Pop volvieron a centrarse en el público adolescente, persiguiendo el éxito continuo con producciones en cierto grado artificiales, que ocupaban las listas de ventas durante el tiempo justo para ser sustituidas por otras nuevas. El trío de productores ingleses Stock, Aitken y Waterman fue el máximo valedor de esta corriente, apadrinando a artistas como Rick Astley, Jason Donovan o Kylie Minogue. En la siguiente década abundarían también los grupos de cantantes y bailarines, como New Kids on The Block o Take That. Al mismo tiempo, el pop fue durante esos años el vehículo para la creación de estrellas mediáticas como Madonna o Michael Jackson , en consonancia con el espíritu frívolo que caracterizó a la sociedad occidental durante esos años.

Sólo los que supieron reinventarse consiguieron mantener cierta relevancia en el show business durante los lustros siguientes. Con la aparición a principios de la década de 1990 del grunge y la música electrónica y pop, el género pop propiamente dicho adquirió la condición de reliquia, aunque ( maquillado bajo producciones heterogéneas ) sigue siendo el punto de partida de muchos compositores actuales.

## Rock

No obstante los evidentes vínculos que posee con las más profundas esencias de la música de la comunidad afroamericana, la musica rock & roll es en realidad el resultado, tras una larga síntesis que se inicia con el siglo, de la adaptación de estas fuentes a una concepción y una estética “blanca”, lo que generó las primeras actitudes públicas de rechazo, algunas de ellas muy conflictivas. De esta forma fue considerado por compositores de blues, tales como Muddy Waters, “un hijo del rythm and blues al que llamaron rock”, siendo numerosas organizaciones racistas estadounidenses las que afirmaban con desprecio que este nuevo tipo de música “rebajaba al hombre blanco a la categoría de negro”. La variante, en cualquier caso, de acuerdo con estas interpretaciones, radicaría en el hecho de que la base del rock & roll, el viejo blues, estaba revisada, para mayor gloria de la raza blanca, a través de inyecciones de ritmo tan pronto dulce, como de andanadas de sonido estridente, rápido y energético.

Durante mucho tiempo este nuevo estilo de música fue llamado como una síntesis ligera de blues y R&B, y hallaría su ídolo fundacional a mediados de la década de 1950 en la figura de Elvis Presley, aún cuando desde la primera grabación de blues registrada en la historia ( “Crazy Blues”, un tema cantado en 1920 por Mammie Smith ) hubieran pasado casi tres decenios.

Esta vía de ”desencuentro”, conforme con sus raíces negras, la seguirían cantantes como Bessie Smith ( llamada la Emperatriz del Blues ), Billie Holiday, Little Richard y, ya en la esfera de la música de finales de la década de 1960 cantada por mujeres blancas, Janis Joplin. Para entonces, por los espasmódicos bailes que suscitaba en grandes masas de

jóvenes y sus ritmos trepidantes, el rock se había transformado en algo más que música, era Rock and Roll , energía en estado puro.

## Salsa

Los orígenes de la música popular suelen ser un asunto de disputa y, en el caso de la música salsa, esa afirmación genérica se cumple. Varios pueblos y lugares reclaman para sí el honor de haber alumbrado este género musical hispano que ha logrado conquistar una buena parte del mundo. En realidad los principios de este género musical tienen raíces largas y profundas que alcanzan Cuba, Puerto Rico, Venezuela, Colombia y cuyo tronco aflora en Nueva York y Miami.

La palabra “salsa” puede llegar a provenir de cierta reivindicación cubana, basándose en que el cubano Ignacio Piñeiro interpretaba en 1933 una canción, “Échale Salsita”, que además de introducir la trompeta por primera vez en el son, sirvió para que ese culinario elemento se colara en la músicaailable caribeña. No se encuentra ninguna continuidad entre esa aparición inicial y la utilización de la palabra “salsa” como denominación de un determinado género. En realidad, palabras como salsa, azúcar, sabor, pese a ser nombres, se han utilizado en la música caribeña a modo de interjección, como elemento gramatical destinado a poner un acento de expresividad no directamente relacionado con su estricto significado. Dado este uso, surge el nombramiento de una música de profundos orígenes cubanos transformada y reelaborada en los barrios hispanos de Nueva York. Se puede decir que el término Salsa fue afirmado como la denominación de un género gracias a un hallazgo comercial de la discográfica Fania Records al principio de los años 70 en Nueva York. De forma desmitificadora decía Tito Puente en una entrevista que ”salsa es lo que yo como con mis espaguetis pero gracias a esa pequeña palabra que no significa nada, todos nosotros hemos sido capaces de encontrar un modus vivendi durante las últimas décadas”.

## 2.3 El audio digital

Este término hace referencia a la tecnología que graba, almacena y reproduce sonido de una señal de audio codificada de manera digital en lugar de una forma analógica [16].

Los sistemas de audio digital pueden incluir compresión, almacenamiento, procesamiento y transmisión de sus componentes.

El audio digital es muy usado en grabación, manipulación, producción masiva y distribución del sonido.

Dado que el sonido es una onda de presión se requiere un transductor de presión (un micrófono) que convierte las ondas de presión de aire (ondas sonoras) en señales eléctricas (señales analógicas).

La conversión contraria se realiza mediante un altavoz, que convierte las señales eléctricas en ondas de presión de aire.

Un sólo micrófono puede captar adecuadamente todo el rango audible de frecuencias, en cambio para reproducir fidedignamente ese mismo rango de frecuencias suelen requerirse dos altavoces (de agudos y graves) o más.

Una señal de audio se puede caracterizar, sonoramente por su valor de pico, rango dinámico, potencia, ruido, ancho de banda, frecuencia fundamental, armónicos, distorsión armónica, etc.

Así, por ejemplo, una señal que represente voz humana (señal vocal) no suele tener información relevante más allá de los 10 kHz, y de hecho en telefonía fija se toman sólo los primeros 3.8 kHz. Con 2 kHz basta para que la voz sea comprensible, pero no para recono-

cer al hablante.

## 2.4 Digitalización del sonido

### 2.4.1 Conceptos

Las variaciones que produce el sonido se denominan comúnmente formas de onda. La duración de un motivo recurrente de la forma de onda es el periodo  $T$ . El número de veces que dicho periodo se repite en un segundo (la inversa del periodo) nos da su frecuencia fundamental  $f$ ,  $f = \frac{1}{T}$ . La frecuencia fundamental de una onda determina su altura. Si el sonido solo contiene una frecuencia, decimos que el sonido es sinusoidal puro. Los sonidos naturales contienen diversas frecuencias, además de la fundamental, que forman un sonido complejo y dan lugar a la percepción del timbre. A las frecuencias adicionales se les denomina armónicos o parciales, según sea su frecuencia múltiplo o no de la frecuencia fundamental [9].

### 2.4.2 Digitalización

Digitalizar el sonido consiste en convertir en código binario los cuatro parámetros del sonido, para cada segmento de tiempo hay que dar información exacta respecto a altura, volumen, duración y timbre. Dicha tarea la realizan los conversores analógico-digital (AD) y digital-analógico.

Puesto que el sonido se transmite por ondas analógicas, los sistemas de cómputo convierten la señal en una señal digital, a este proceso se le denomina digitalización del sonido.

La digitalización del sonido consiste en tomar muestras de la señal analógica cada instante de tiempo, cuanta más veces se mida la señal sonora y mayor sea el rango para medir la intensidad, mejor intensidad tendrá el sonido digitalizado y más se parecerá al sonido real.

## 2.5 Características del sonido

El sonido posee diversas características que lo permiten diferenciar y distinguir. Algunas de ellas son el dominio de la frecuencia y otras en el dominio del tiempo.

### 2.5.1 En el dominio del tiempo

#### Volumen

El volumen es la intensidad con la que se percibe un sonido en determinado tiempo (un segundo). Es lo inverso al silencio, es decir que tan ruidoso es un sonido. Es necesario normalizar el volumen con respecto a las segmentos con mayor volumen para eliminar las condiciones de los dispositivos que influyen.

Para calcular el aproximado del volumen de un segmento de audio se utiliza la media cuadrática (RMS Root Mean Square) como se describe en la siguiente fórmula donde el volumen se representa por  $V(n)$ .

$$V(n) = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} S_n^2(i)} \quad (2.1)$$

donde  $N$  es la longitud del segmento de la señal del que se desea obtener el volumen y  $S_n(i)$  es la  $i$ ésima muestra del segmento [23].

#### Tono

El tono es la frecuencia fundamental de la forma de onda y es un parámetro importante en el análisis del sonido. El tono de un sonido generalmente se describe como que tan alto o bajo (grave o agudo) es percibido. Es la frecuencia fundamental de la onda de sonido, el rango de tono de la voz humana se encuentra entre los 50 y 450 Hz, aunque para la música el rango es mucho más amplio.

Existen diversos métodos para obtener una aproximación del tono, dependiendo del nivel

de precisión que se requiera. Para extraer información del tono se puede hacer uso del análisis en el tiempo o en la frecuencia. Aunque para hacer un cálculo computarizado es mejor obtener el tono en el dominio del tiempo ya que en el dominio de la frecuencia se consumen más recursos.

⇒ Estimación temporal en un tiempo corto.

La estimación temporal en un tiempo corto se encuentra dada por la función  $R_n(l)$ .

$$R_n(l) = \sum_{i=0}^{N-l-1} S_n(i)S_n(i+l) \quad (2.2)$$

⇒ Función promedio de diferenciación de magnitud (AMDF Average Magnitud Difference Fuction). Esta función se representa a continuación por  $A_n(l)$ .

$$A_n(l) = \sum_{i=0}^{N-l-1} |S_n(i+l) - S_n(i)| \quad (2.3)$$

Donde  $N$  representa la longitud del segmento de la señal y  $S_n(i)$  la  $i$ ésima muestra del segmento [23].

Calificación de cruce por cero (ZRC Zero Crossing Rate)

La calificación de cruce por cero también es una característica en el dominio del tiempo, en el que se cuenta el número de veces que la onda cruza por cero. Formalmente se define con la siguiente fórmula.

$$Z(n) = \frac{1}{2} \sum_{i=1}^{N-1} |\text{sign}(S_n(i)) - \text{sign}(S_n(i-1))| \frac{f_s}{N} \quad (2.4)$$

Donde  $Z(n)$  es la calificación de cruce por cero,  $N$  es la longitud del segmento de señal,  $S_n(i)$  la  $i$ ésima muestra del segmento y  $f_s$  la frecuencia de muestreo [23].

## 2.5.2 En el dominio de la frecuencia

### Centro del espectro

Describe el centro de gravedad del espectro y se describe mediante la siguiente  $C_i$

$$C_i = \frac{\sum_{k=1}^{N/2} f(k) |X_i(k)|}{\sum_{k=1}^{N/2} |X_i(k)|} \quad (2.5)$$

donde  $N$  es el número total de puntos en la FFT,  $X_i(k)$  es el poder o amplitud del  $k$ ésimo punto del  $i$ ésimo frame,  $f(k)$  es la frecuencia correspondiente al punto de la FFT [3].

### Roll-off

Es una medida de la forma espectral. Se define como la frecuencia por debajo del cual se concentra  $C$  porcentaje de las magnitudes acumuladas del espectro, donde  $C$  es un coeficiente empírico. Esta medida de la forma del espectro, es usada para distinguir si existe o no el sonido de una voz [3].

$$\sum_{k=1}^K X_i(k) \leq C \cdot \sum_{k=1}^K X_i(k) \quad (2.6)$$

### Flujo

Es usado para describir la tasa de cambio del espectro. Se calcula con la diferencia total entre la magnitud de los puntos de FFT (Fast Fourier Transform) de un fragmento con el fragmento siguiente. El flujo  $F_i$  se puede calcular de la siguiente manera

$$F_i = \sum_{k=1}^{N/2} (|X_i(k)| - |X_{i-1}(k)|)^2 \quad (2.7)$$

donde  $N$  es la longitud del  $i$ -ésimo segmento de señal que se analiza y  $X_i(k)$  es la amplitud o el poder de del  $k$ -ésimo punto de la FFT [3].

### 2.5.3 El Cepstrum

El cepstrum puede ser visto como una información del ritmo de cambio de las diferentes bandas de un espectro. Fue originalmente desarrollada para caracterizar los ecos de las ondas tectónicas, que provenían de terremotos y explosiones. También se utilizaba para analizar las señales captadas por un radar. El espectro de potencia  $|F\{\log(|F\{f(t)\}|^2)}|^2$  es usado como un vector de características para representar la voz humana y las señales de música. Para este tipo de aplicaciones el cepstrum es transformado usando la escala melódica. El resultado de dicha transformación es el cepstrum melódico de frecuencia (mel frequency cepstrum coefficient). Los MFCC son usados en la identificación de voz, detección de tono, etc. [15]

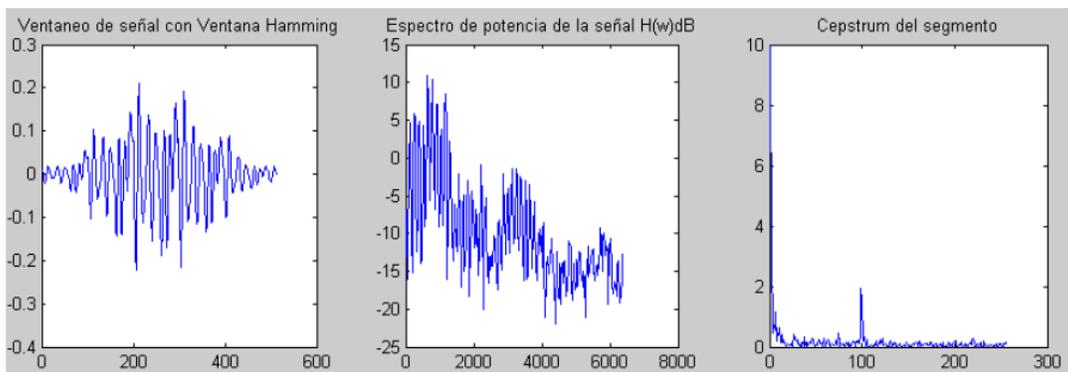
El cepstrum de una señal es el resultado de calcular la transformada de Fourier (FT, Fourier Transform) del espectro de la señal estudiada en escala logarítmica (db). El nombre cepstrum deriva de invertir las primeras cuatro letras de spectrum. El cepstrum es complejo y, por tanto tiene su parte real y su parte imaginaria. Se puede definir de las siguientes maneras:

- ⇒ Verbalmente: El cepstrum de una señal es la transformada de Fourier del logaritmo del espectro de la señal estudiada. A veces es llamado espectro del espectro.
- ⇒ Algorítmicamente: Señal  $\rightarrow$  FT  $\rightarrow$  Valor absoluto  $\rightarrow$  log  $\rightarrow$  Ajuste de la fase instantánea  $\rightarrow$  FT  $\rightarrow$  Cepstrum.

La variable independiente de un gráfico de cepstrum se llama quefrequency. La quefrequency es una medida del tiempo, pero no en el sentido del dominio temporal. Por ejemplo si la velocidad de muestreo de una señal de audio es de 44.100 Hz y hay un gran pico en el cepstrum en la quefrequency de 100 muestras, este pico indica la presencia de un pitch (frecuencia fundamental percibida) a  $44100/100 = 441$  Hz. Este pico que aparece en el cepstrum indica

entonces el periodo en el que están los armónicos del espectro. Nótese que una sola onda sinusoidal no se debiera usar como test para el cepstrum ya que esta corriente no contiene armónicos. En su lugar, debiera usar una señal de test que sí contenga armónicos, como la suma de dos senos en la cual el segundo tenga una frecuencia múltiplo de la primera (armónico). [15]

En la figura 2.5.1 se ilustra la obtención del cepstrum de un fragmento de señal en la frecuencia del tiempo a la que se le aplica una ventana de Hamming.



**Imagen 2.5.1:** Cepstrum de un fragmento de señal. [8]

Coeficientes cepstrales en la frecuencia de Mel (MFCC, Mel-frequency cepstral coefficients)

Son coeficientes que en conjunto conforman un cepstrum en la frecuencia de Mel (MFC, Mel-frequency cepstrum). Se derivan de un tipo de representación cepstral de un segmento de audio. La diferencia entre el cepstrum y el cepstrum en la frecuencia de Mel MFC es que en el MFC, las bandas de frecuencias están igualmente espaciados en la escala de mel, que se aproxima a la respuesta del sistema auditivo humano más estrechamente que las bandas de frecuencia linealmente espaciados utilizados en el cepstrum normal. Esta deformación de frecuencia puede permitir una mejor representación del sonido, por ejemplo, en la compresión de audio. [17]

Los MFCC se derivan generalmente del siguiente modo: [5] [12]

- ⇒ Tomar la transformada de Fourier
- ⇒ Mapear los poderes del espectro obtenidos anteriormente en la escala de Mel, utilizando la superposición de ventanas triangulares.
- ⇒ Tomar los registros de las potencias en cada una de las frecuencias de Mel.
- ⇒ Tomar la transformada discreta del coseno de la lista de potencias de registro de Mel, como si fuera una señal.
- ⇒ Los coeficientes MFCC son las amplitudes del espectro resultante.

## 2.6 Archivos de sonido

En la actualidad se pueden distinguir dos tipos de archivos de sonido, con y sin pérdida, esto es debido a que se usan algoritmos de compresión, generalmente usados en información analógica al momento de ser digitalizada, también se aplican algoritmos de compresión en imágenes y video.

### 2.6.1 Archivos de sonido con pérdida

Estos archivos usan algoritmos de compresión con pérdida, esto significa que utilizan menos cantidad de información. Estos archivos tienen la ventaja de que ocupan un menor espacio de almacenamiento pues los datos digitalizados son menos, haciendo así imposible reconstruir el archivo original.

- ⇒ **MP3** o MPEG-1 Audio Layer 3: Formato de audio estándar cuya pérdida de información no es percibida por el oído humano. Este formato ha llegado a ser un estándar, por lo que es el más usado en reproductores portátiles teniendo una compatibilidad con los medios garantizada. Su frecuencia de muestreo ronda entre los rangos de 16 Hz y los 48 KHz.

- ⇒ **ACC** o Advanced Audio Coding es un formato de audio digital estándar como extensión de MPEG-2, ofreciendo mayor calidad que el formato MP3 aunque su principio de compresión sea el mismo, esto es por el rango de frecuencias que se encuentra entre los 8 Hz hasta los 96 KHz. Aunque ofrece mayor calidad de audio y mejor compresión que el MP3 no tiene la misma popularidad por lo tanto su uso y compatibilidad es menor.
- ⇒ **Ogg** formato comprimido por el códec Vorbis que es un códec de audio libre que ofrece una muy buena calidad de sonido ya que su muestreo va desde los 8 KHz hasta los 48 KHz. También ofrece una mejor calidad y compresión que el MP3 pero aún no es muy utilizado debido a su poca popularidad.
- ⇒ **Real Audio** o RM es un formato muy usado para las transmisiones por Internet en tiempo real, pues su codificación se adapta a la recepción del usuario reduciendo la frecuencia del muestreo hasta que el usuario pueda recibir paquetes de audio completos que se almacenan en una carpeta temporal hasta que sean requeridos.
- ⇒ **WMA** o Windows Media Audio está desarrollado con fines comerciales para el reproductor de Windows Media Player de Windows. Es un formato de compresión de audio con pérdida aunque también existe este formato con compresión sin pérdida.

## 2.6.2 Archivos de sonido sin pérdida

Los archivos sin pérdida son aquellos que usando o no usando métodos de compresión, representan la información sin desprender información original, por lo tanto hace posible reconstruir de manera exacta la información original.

- ⇒ **AIFF** o Audio Interchange File format que significa Formato de Archivo de Intercambio de Audio, los archivos en este formato poseen una modulación por impulsos codificados o PCM. Es usado a nivel profesional en equipos Apple y Amiga, teniendo como desventaja la gran cantidad de espacio en disco que utiliza aproximadamente 10 MB para un minuto. La frecuencia de muestreo es de 44.1 KHz y 16 bits. Sus extensiones son .aif, .aiff y .aifc. esta última es una variante que si posee compresión.

- ⇒ **FLAC** o Free Lossless Audio Codec es un método sin pérdida pero que aún así se consigue reducir el tamaño del archivo original entre la mitad y tres cuartos haciendo así posible recuperar el archivo original a partir de un formato flac. Su rango de muestreo es aproximadamente desde 1 hasta 65535 KHz en incrementos de 1 Hz. Este formato es muy usado en la venta de música por Internet.
- ⇒ **WAV** o wave (Waveform Audio Format) es un formato digital sin compresión y de el se habla más adelante.
- ⇒ **MIDI** Interface Digital para Instrumentos Musicales, es considerado el estandar para la música electrónica ya que se puede utilizar con sintetizadores musicales o tarjetas de sonido. Su extensión es .midi o .mid.

## 2.7 Archivo Wav

Es un archivo de audio de forma de onda, también conocido como un archivo de onda, o simplemente WAV. Es un tipo común de archivo de sonido. Microsoft e IBM introdujo el archivo wav en 1991 para su uso en el sistema de operación de Microsoft Windows 3.1 (OS).

El formato WAV se basa en el formato de intercambio de archivos de recursos (RIFF), que almacena los archivos de audio en "trozos" y "sub-trozos". RIFF es a su vez basado en el formato de archivo de intercambio anterior (IFF), creado por Electronic Arts en 1985 para su uso en juegos electrónicos.

En el archivo WAV hay dos cosas muy importantes a su favor. En primer lugar, al digitalizar suena 100% fiel a la fuente original, ya que es un formato sin pérdidas. "Lossless" significa que el formato de archivo WAV no compromete la calidad de audio incluso cuando se tienen datos comprimidos. En segundo lugar, el archivo WAV es muy fácil de editar y manipular con software.

Mientras que el formato de archivo WAV era ideal para efectos de sonido, tenía una desventaja cuando se trata de archivos de música. Una canción de cuatro minutos puede fácilmente consumir más de 35 megabytes (MB) de espacio cuando se guarda como un archivo WAV. Aunque el costo de los discos duros ha disminuido en los últimos años, el formato de archivo wav aún era demasiado grande para reproductores portátiles con tarjetas de memoria limitada. Además, los archivos wav no son el formato más práctico para transferir en línea, especialmente en conexiones lentas de acceso telefónico.

En cambio, el formato comprimido MP3, subió al escenario audio. Las canciones guardadas en el formato de archivos de audio MP3 produjo una mera fracción del tamaño de los archivos wav. El formato MP3 es un formato con pérdida. Mientras que el formato MP3 es una buena opción para reproductores portátiles, mucha gente prefirió almacenar sus principales bibliotecas digitales en formato de archivo wav. De este modo, el archivo wav se podría utilizar como un maestro para crear otros tipos de archivos de audio (incluyendo MP3).

Hoy en día el formato de archivo wav sigue siendo ampliamente utilizado para los archivos de música en un formato sin pérdidas, donde el espacio no es un problema. Algunos reproductores de CD y DVD también pueden leer los archivos wav copiados directamente en un disco compacto.

#### 2.7.1 Limitantes del formato Wav

El formato WAV se limita a archivos que son menos de 4 GB, debido a su uso de un entero sin signo de 32-bits para registrar el tamaño del archivo de cabecera.

A diferencia de formatos como FLAC, WAV no suelen tener campos de información, por ejemplo, en el caso de una canción, título, artista, álbum, año, etc

## 2.8 Redes Neuronales

La investigación en el campo de las redes neuronales ha estado atrayendo cada vez más atención en los últimos años. Desde 1943, cuando Warren McCulloch y Walter Pitts presentaron el primer modelo de neuronas artificiales, nuevos y más sofisticados se han hecho propuestas de década en década. El análisis matemático tiene resolver algunos de los misterios que plantean los nuevos modelos, pero ha dejado muchas preguntas abiertas para futuras investigaciones. No hace falta decir, el estudio de las neuronas, su interconexiones, y su papel como bloques de construcción del cerebro elementales es uno de los campos de investigación más dinámicos e importantes de la biología moderna. nosotros puede ilustrar la importancia de este esfuerzo por señalar que entre 1901 y 1991, aproximadamente el diez por ciento de los Premios Nobel de Fisiología y Medicina fueron otorgados a científicos que han contribuido a la comprensión de el cerebro. No es una exageración decir que hemos aprendido más sobre del sistema nervioso en los últimos cincuenta años que nunca antes.

Las redes neuronales artificiales son un intento de modelar el procesamiento de la información capacidades de los sistemas nerviosos. Así, en primer lugar, debemos tener en cuenta las propiedades esenciales de redes neuronales biológicas desde el punto de vista de la información procesar. Esto nos permitirá diseñar modelos abstractos de artificial redes neuronales, que luego pueden ser simulados y analizados.

Aunque los modelos que se han propuesto para explicar la estructura del cerebro y el sistema nervioso de algunos animales son diferentes en muchos aspectos, existe un consenso general de que la esencia de la operación de conjuntos neuronales es "el control a través de la comunicación".

Los sistemas se componen de miles o millones de células interconectadas. cada una de ellas es una disposición muy compleja que trata las señales entrantes de muchas maneras diferentes. Sin embargo, las neuronas son más bien lentos en comparación con los puertas electrónicas lógicas. Estos pueden lograr tiempos de conmutación de unos pocos nanose-

gundos, mientras que las neuronas necesitan varios milisegundos para reaccionar a un estímulo. sin embargo el cerebro es capaz de resolver problemas que ninguna computadora digital puede aún eficientemente tratar.

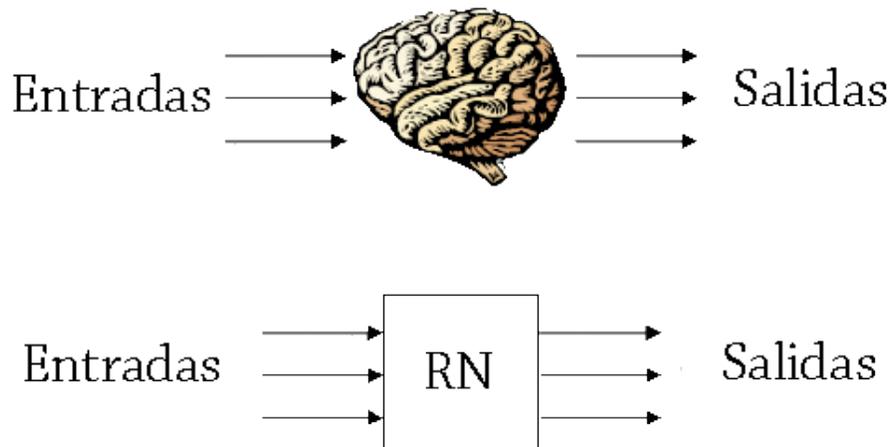
### 2.8.1 Aprendizaje No Supervisado

Este se utiliza cuando, para una determinada entrada, la salida numérica exacta de una red debe producir es desconocida. Supongamos, por ejemplo, que algunos puntos en el espacio bidimensional se clasifican en tres grupos. Para esta tarea se puede utilizar una red clasificador con tres líneas de salida, uno para cada clase. Cada una de las tres unidades de computación en la salida debe especializarse por el disparo sólo para las entradas correspondientes a los elementos de cada grupo. Si una unidad de incendios, los otros deben permanecer en silencio. En este caso, no sabemos a priori que la unidad se va a especializar en que cluster. En general, no sé ni cuántos grupos bien definidos están presentes. Puesto que ningún "maestro" está disponible, la red debe organizarse con el fin de ser capaz de asociar con grupos de unidades.

### 2.8.2 Aprendizaje Supervisado

El aprendizaje supervisado se divide en los métodos que utilizan refuerzo o corrección de errores. Aprendizaje por refuerzo se utiliza cuando después de cada presentación de un ejemplo de entrada-salida sólo se sabe si la red produce el resultado deseado o no. Las ponderaciones se actualizan sobre la base de esta información (es decir, los valores booleano verdadero o falso) para que sólo el vector de entrada se puede utilizar para la corrección de peso. En el aprendizaje con corrección de errores, la magnitud del error, junto con el vector de entrada, determina la magnitud de las correcciones a los pesos, y en muchos casos se intenta eliminar el error en un paso de corrección único.

La imagen 2.8.1 muestra un diagrama de bloques básico de las redes neuronales.



**Imagen 2.8.1:** Diagrama de bloque de una red neuronal

## 2.9 Paradigma clasificatorio K-NN (K - Nearest Neighbour)

La idea básica sobre la que se fundamenta este paradigma es que un nuevo caso se va a clasificar en la clase más frecuente a la que pertenecen sus K vecinos más cercanos. El paradigma se fundamenta por lo tanto en una idea muy simple e intuitiva, lo que unido a su fácil implementación hace que sea un paradigma clasificatorio muy extendido. [1]

### 2.9.1 Algoritmo K-NN básico [1]

En la tabla 2.9.1 se muestra la notación a utilizar.

		$X_1$	...	$X_j$	...	$X_n$	$C$
$(x_1, c_1)$	1	$x_{11}$	...	$x_{1j}$	...	$x_{1n}$	$C_1$
	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$(x_i, c_i)$	i	$x_{i1}$	...	$x_{ij}$	...	$x_{in}$	$C_i$
	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$(x_N, c_N)$	N	$x_{N1}$	...	$x_{Nj}$	...	$x_{Nn}$	$C_N$
$x$	N + 1	$x_{N+1,1}$	...	$x_{N+1,j}$	...	$x_{N+1,n}$	?

**Tabla 2.9.1:** Notación para el paradigma K-NN.

⇒  $D$  indica un fichero de  $N$  casos, cada uno de los cuales está caracterizado por  $n$  variable predictorias,  $X_1, \dots, X_n$  y una variable a predecir, la clase  $C$ .

⇒ Los  $N$  casos se denotan por:

$$(x_1, c_1), \dots, (x_N, c_N) \text{ donde:}$$

$$x_i = (x_{i,1} \dots x_{i,n}) \text{ para todo } i = 1, \dots, N$$

$$c_i \in \{c^1, \dots, c^m\} \text{ para todo } i = 1, \dots, N$$

$c^1, \dots, c^m$  denotan los  $m$  posibles valores de la variable clase  $C$ .

⇒ El nuevo caso que se pretende clasificar se denota por  $x = (x_1, \dots, x_n)$ .

En la figura 2.9.1 se presenta un pseudocódigo para el clasificador K-NN básico. Tal y como puede observarse en la misma figura, se calculan las distancias de todos los casos ya clasificados al nuevo caso,  $x$ , que se pretenda clasificar. Una vez seleccionados los  $K$  casos ya clasificados,  $D_x^k$  más cercanos al nuevo caso,  $x$ , a éste se le asignará la clase (valor de la variable  $C$ ) más frecuente de entre los  $k$  objetos,  $D_x^k$ .

COMIENZO

Entrada:  $D = \{(x_1, c_1), \dots, (x_N, c_N)\}$

$x = (x_1, \dots, x_n)$  nuevo caso a clasificar

PARA todo objeto ya clasificado  $(x_i, c_i)$

calcular  $d_i = d(x_i, x)$

Ordenar  $d_i (i = 1, \dots, N)$  en orden ascendente

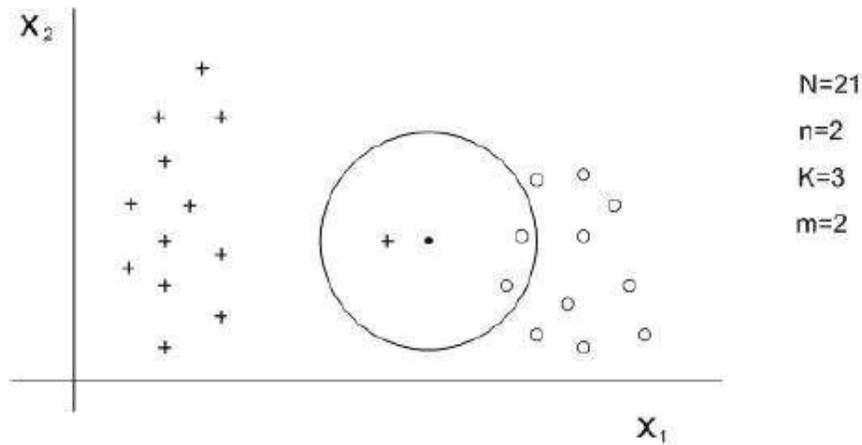
Quedarnos con los  $K$  casos  $D_x^K$  ya clasificados más cercanos a  $x$

Asignar a  $x$  la clase más frecuente en  $D_x^K$

FIN

**Imagen 2.9.1:** Pseudocódigo para el clasificador K-NN

En la figura 2.9.2 se muestra de manera gráfica un ejemplo, en el que se muestran 24 casos ya clasificados en dos posibles valores ( $m = 2$ ). Las variables predictorias son  $X_1$  y  $X_2$ , y se ha seleccionado  $K = 3$ . De los 3 casos ya clasificados que se encuentran más cercanos al nuevo caso a clasificar,  $x$  (representado por  $\bullet$ ), dos de ellos pertenecen a la clase  $\circ$ , por lo tanto el clasificador 3-NN predice la clase  $\circ$  para el nuevo caso. Nótese que el caso más cercano a  $x$  pertenece a la clase  $+$ . es decir, que si hubiésemos utilizado un clasificador 1-NN,  $x$  se hubiese asignado a  $+$ .



**Imagen 2.9.2:** Ejemplo de aplicación del algoritmo K-NN básico

# 3

## Problemática

### 3.1 Descripción de la problemática

Actualmente la mayoría de las aplicaciones existentes y disponibles en el mercado clasifican los archivos por metadatos básicamente mediante una etiqueta de género que se hace manualmente. También existen trabajos de investigación en diferentes países en los que se realizan clasificaciones de archivos audio mediante redes neuronales o utilizando métodos dinámicos.

La manera en la que se realiza la clasificación de la música en cuanto a su contenido musical suele ser subjetiva de acuerdo al criterio del evaluador. Por lo tanto resulta engorroso y tardado que el usuario o evaluador escuche uno a uno sus archivos digitales de audio para así lograr clasificar el audio del archivo en algún género que de acuerdo a su criterio sea el apropiado. Actualmente no existen aplicaciones que logren hacer una clasificación de

manera automática a archivos de audio digitales mediante el análisis del audio en géneros de acuerdo a la similitud de patrones cuantitativos.

Nuestra herramienta propone una clasificación de los archivos de audio por medio de un método de aprendizaje supervisado, que se entrenará inicialmente con 7 conjuntos de canciones, cada conjunto corresponde a los géneros rock, pop, jazz, electrónica, tropical, clásica y hip hop, cabe destacar que estos conjuntos se conformaron mediante el criterio de nosotros como desarrolladores.

Las ventajas que presenta nuestra solución frente a otros trabajos de investigación es que se ampliará el número de géneros de clasificación, en la mayoría de las investigaciones previas se proponen de 3 a 4. Otra ventaja notable es la capacidad de configuración de los géneros propuestos desde el inicio y la oportunidad de crear nuevos géneros para clasificar de acuerdo al criterio del usuario final.

Una comparativa de las características de sistemas existentes con el sistema propuesto se muestra en la tabla 3.1.1

SOFTWARE	CARACTERÍSTICAS	PRECIO EN EL MERCADO
Clasificador de música por genero utilizando redes neuronales artificiales	Se usa Backpropagation, perceptron multicapa, solo archivos .wav y los clasifica en generos Salsa Merengue Vallenato	Proyecto de la Pontificia Universidad Javeriana, Facultad de Ingeniería, Bogotá.
Diseño de un clasificador de géneros musicales basado en modelos dinámicos	Usa métodos dinámicos, escalado y supervisado con los Coeficientes de Mel y Modelos Ocultos de Markov	Proyecto de la Universidad Carlos III de Madrid, Escuela Politécnica Superior, Ingeniería de Telecomunicación.
Representación topológica de las características de la señal digitalizada de la música para la generación automática de listas de reproducción	Genera listas de reproducción de música mediante la combinación de diferentes parámetros	Proyecto de la Universidad Técnica Federico Santa María, Departamento de Informática, Valparaíso, Chile.
Score a Hit	Aplicación Web: Analiza archivos de audio y compara con otros archivos de audio de las listas Billboard y muestra si pudiera ser un éxito o no.	Gratis, Univerisdad de Bristol
Genius	Aplicación para iphone, ipad, ipod. Genera listas de reproducción a partir de una pista musical seleccionada, la lista propuesta se basa en las listas de reproducción que poseen los demás usuarios de esta herramienta.	Gratis en Itunes, Apple
Solución propuesta: Clasificador Unimodal de Archivos de Audio	Funcionamiento con archivos de audio en formato WAV dependiendo del género (rock, pop, jazz, electrónica, tropical, clásica, hip-hop u otro). Esto se logrará mediante el uso de algoritmos de clasificación y la extracción de características del la señal de audio contenida en el archiv musical.	Proyecto de investigación

**Tabla 3.1.1:** Tabla de sistemas similares y solución propuesta

# 4

## Análisis

### 4.1 Estudio de factibilidad

#### 4.1.1 Factibilidad Operacional

El propósito de este estudio operacional es realizar un análisis que determine si el Clasificador Unimodal de Archivos Musicales se podrá utilizar en la forma en la que se supone.

En la actualidad no existe software comercial que realice una clasificación de archivos de audio musicales por género musical. La mayoría de las soluciones para organizar bibliotecas musicales personales son capaces de organizar el contenido de la biblioteca de acuerdo a las etiquetas de los archivos de audio (metadatos), mismos que son introducidas de forma manual por el usuario u obtenidas por medio de los reproductores multimedia a través de una conexión a Internet. Incluso existen soluciones que son capaces de crear listas de re-

producción acorde a una pista seleccionada con la desventaja de que la selección de pistas se lleva a cabo de acuerdo a la recopilación de información de las listas de reproducción de otros usuarios de la misma solución, teniendo así la desventaja de que no siempre existen listas de reproducción disponibles para cualquier pista de audio que el usuario seleccione y la necesidad de contar con una conexión a Internet.

El propósito primordial de la solución propuesta es que el usuario obtenga el género sus archivos musicales de forma automática de acuerdo a géneros musicales previamente establecidos.

Se presenta el problema de que la mayoría de los usuarios ya se encuentran muy familiarizados con sistemas de reproducción multimedia, complicando así el lograr que el usuario use soluciones ajenas a las que le ofrece su reproductor multimedia tradicional.

Otro gran reto que tenemos es con el formato de archivos de audio, la mayoría de los reproductores comerciales son capaces de manejar hasta más de 3 o 4 formatos de archivos de audio, siendo el objetivo primordial del Clasificador Unimodal de Archivos Musicales lograr la clasificación por géneros sin poner mucha atención en la compatibilidad de formatos de archivos.

También el reto que se ha propuesto es el de lograr etiquetar con metadatos los archivos musicales y con esto se podrá lograr una compatibilidad entre reproductores multimedia, ya que la mayoría de estos reconocen metadatos y a partir de estos que se realice la organización, mas no la clasificación. Así el Clasificador Unimodal de Archivos podrá ser una utilidad de gran ayuda para el usuario. Sin embargo este reto se pospondrá para trabajos futuros, pues se seleccionó el formato WAV para trabajar en esta investigación, mismo que no posee etiquetas de metadatos como por ejemplo el formato MP3 que utiliza las etiquetas ID3.

Se concluye que es factible operacionalmente el desarrollo del Clasificador Unimodal de

Archivos como una aplicación de investigación para la clasificación por género de archivos musicales.

#### 4.1.2 Factibilidad Técnica

El proyecto resulta viable técnicamente ya que en la actualidad el campo de estudio en el análisis de señales ha aumentado día a día, dando lugar a una serie de operaciones, algoritmos y transformadas que son capaces de brindar información de utilidad, se debe prestar especial atención a la manera en la que se implementarán dichas técnicas en lenguajes de programación.

Existen variedad de lenguajes de programación, de los cuales algunos cuentan con módulos o librerías para el tratamiento de señales de audio.

Por otro lado no se ha contemplado el uso de algún hardware externo además de los equipos de cómputo personales.

En conclusión el Clasificador Unimodal de Archivos resulta técnicamente viable.

#### 4.1.3 Factibilidad Económica

En el apartado correspondiente al estudio del hardware y software que se utilizarán durante el desarrollo del presente proyecto se identifica que ya se cuenta con el equipo de hardware que es requerido, ya que es una herramienta indispensable para los integrantes, además de que se usan distribuciones de software libre.

La tabla 4.1.1 muestra una relación de las horas de trabajo que se estima que cada estudiante invierta y un costo aproximado del costo por hora. Cabe destacar que en la tabla no se incluyen las horas de trabajo de los directores ni del supervisor.

Hrs semana	Hrs mes	Costo hrs	Costo mes	Años	Costo integrante	Total x 2 integrantes
15	60	\$60.00	\$ 3,600.00	2	\$86,400.00	\$172,800.00

**Tabla 4.1.1:** Tabla de costo de horas de trabajo.

Es importante mencionar que durante el proceso del desarrollo del presente trabajo ninguno de los estudiantes cobrará por sus servicios, puesto que realiza esta actividad para adquirir el título de Ingeniero en Sistemas Computacionales como un proyecto de investigación.

En la sección de estudio de software y hardware se detallan los costos aproximados de los mismos y se muestran algunas tablas que describen mejor esta información, se dedujo que no hay necesidad de realizar alguna inversión económica para adquirir equipo que nos permita continuar con el presente proyecto de investigación debido a que ya se cuenta con el equipo y el software de licencia libre. Podemos concluir que el proyecto es viable económicamente.

#### 4.1.4 Factibilidad Legal

Al realizar el Clasificador Unimodal de Archivos no se incurre en ninguna infracción de tipo legal, se sujeta a los lineamientos sobre el uso de licencias GNU GPL (General Public License), LPPL (Licencia Pública del Proyecto LaTeX).

## 4.2 Estudio de software y hardware

Para la puesta en marcha del proyecto se hará uso de hardware, software y de archivos musicales.

### 4.2.1 Hardware

En cuanto al hardware, se hará uso de un equipo de cómputo portátil por cada integrante del equipo, ambos portátiles a continuación son mostrados en la tabla siguiente con sus caracte-

terísticas y precios aproximados de acuerdo a cotizaciones realizadas el día de elaboración del documento.

En la Tabla 4.2.1 que se muestra a continuación se detallan las características y costos aproximados de los equipos que van a ser usados durante el presente proyecto.

Equipo	RAM	Disco duro	Procesador	Precio estimado
Laptop Hp Dv6000	4 GB	350 GB	AMD Athlon X2	\$ 4,500.00
Laptop Sony Vaio Vpcek	4 GB	500 GB	Intel Core i3	\$ 10,300.00
Costo total hardware	-	-	-	\$ 14,800.00

**Tabla 4.2.1:** Tabla de hardware requerido.

#### 4.2.2 Software

En la tabla 4.2.2 se enlistan sistema operativo, interfaz gráfica de desarrollo, lenguajes de programación entre otros elementos de software que utilizarán los equipos de hardware mencionados anteriormente (4.2.1).

Software	Versión	Distribuidor	Unidades	Precio
S.O. Ubuntu	10.04	Comunidad Ubuntu	1	libre
S.O. Ubuntu	12.04	Comunidad Ubuntu	1	libre
Java	6	Oracle	2	libre
Python	2	Python Software Foundation	2	libre
NetBeans IDE	6.9	NetBeans Community	2	libre
Notepad++	6.1	Notepad++project	2	libre
LaTeX	2	Proyecto Latex	2	libre
Adobe Reader	9.5	Adobe	2	libre
gedit	2.2	GNOME	2	libre

**Tabla 4.2.2:** Tabla de software requerido.

### 4.2.3 Archivos de audio

Para poder entrenar los algoritmos de clasificación, hacer pruebas de funcionalidad y pruebas de implementación debemos de contar con repositorios de archivos de pistas musicales, estos repositorios son proporcionados por los miembros del equipo de trabajo, quienes los han ido recopilando a través del tiempo en su colección personal.

## 4.3 Análisis de requerimientos

### 4.3.1 Requerimientos funcionales

- ⇒ Ser capaz de leer archivos de audio WAV.
- ⇒ Analizar características de señales digitales de los archivos de audio.
- ⇒ Clasificar archivos de audio de acuerdo a alguna taxonomía.

### 4.3.2 Requerimientos no funcionales

- ⇒ Almacenar características en documentos.
- ⇒ Almacenar los documentos de las características en directorios diferentes.
- ⇒ Contar con una interfaz gráfica amigable al usuario.
- ⇒ Debera calcular las características rápidamente.
- ⇒ La calidad de los archivos WAV debera ser de 16,000 Hz.
- ⇒ La profundidad de los archivos WAV debe ser de 16 bits.
- ⇒ Desarrollado e implementado para plataformas de software libre.

## 4.4 Análisis de riesgos

### 4.4.1 Objetivo del análisis de riesgos

El objetivo del análisis es lograr identificar los riesgos, categorizarlos y si es posible tener un plan en caso de contingencia para minorizar los errores, daños o secuelas que estos provoquen.

### 4.4.2 Identificación de riesgos

En la tabla 4.4.1 se muestran los riesgos identificados, acompañados de un identificador interno, una estimación del porcentaje de probabilidad de que ocurran y una escala subjetiva (bajo, medio, alto) según la severidad que representa el riesgo.

Riesgo	ID	Probabilidad	Severidad
Capacidad de procesamiento limitada para el funcionamiento del sistema	R1	20%	alta
Avería en algún equipo de cómputo de los integrantes que no permita continuar desarrollando	R2	15%	medio
Cantidad insuficiente de registros para pruebas y realizar el primer aprendizaje	R3	20%	baja
Obtención de características inútiles para la clasificación por género	R4	35%	alto
Tiempo insuficiente para investigación de métodos para la extracción de características de señales de audio	R5	30%	alto
Tiempo insuficiente para la implementación de algoritmos de análisis de señales	R6	30%	alto
Suspensión temporal de actividades de investigación o desarrollo de algún integrante del equipo	R7	25%	medio
Descerción de alguno de los integrantes al proyecto	R8	10%	alto

**Tabla 4.4.1:** Tabla de riesgos.

#### 4.4.3 Plan de riesgo

El objetivo de realizar este plan de riesgos es tener claros y al alcance las acciones que se tomarán al momento de presentarse alguno de los riesgos para minorizar la severidad y contingencia que pueda causar.

La tabla 4.4.2 muestra los planes de acción que se seguirán si se llega a presentar un riesgo, los riesgos se muestran identificados por su ID.

ID Riesgo	Plan de acción
R1	Al momento de estar realizando las pruebas se estará monitoreando el tiempo y el desempeño del equipo para descartar o modificar los procesos que consuman demasiados recursos del ordenador.
R2	El integrante del equipo al que se le llegue a presentar la avería deberá hacer lo posible para reemplazar o reparar su equipo de ser posible en menos de una semana.
R3	Emprender la tarea de recolectar más pistas musicales a través de aportaciones voluntarias de compañeros externos al equipo o de descargas de Internet.
R4	Continuar con la investigación y desarrollo de métodos del análisis de señales digitales.
R5	Empeñarnos en una configuración, combinación de los métodos ya obtenidos hasta lograr el máximo desempeño posible.
R6	El integrante en cuestión realizará o invertirá más tiempo de lo previsto al trabajo después de su ausencia y por lo mientras el compañero presente observará la carga de trabajo del ausente cuando esto aplique.
R7	El integrante realizará las tareas del compañero ausente si es que se permite normativamente continuar con el proyecto.

**Tabla 4.4.2:** Tabla con plan de riesgos.

# 5

## Productos o resultados esperados

Al finalizar el proyecto se hará entrega de los siguientes elementos:

- ⇒ Aplicación de escritorio para la clasificación de archivos de audio en formato WAV.
- ⇒ Reporte técnico.

### 5.1 Metodología

Se ha decidido usar la metodología del Prototipo Evolutivo por las siguientes razones.

- ⇒ Se ha planeado desarrollar cada módulo que integra el sistema como un prototipo.
- ⇒ La especificación del sistema se puede desarrollar de forma creciente.
- ⇒ Un prototipo puede ayudar en la obtención y validación de los requerimientos del sistema.

⇒ Se pueden explorar soluciones particulares para la solución de la tarea de cada módulo.

⇒ Mejora en la calidad del diseño.

La figura 5.1.1 se muestra el modelo de Prototipo Evolutivo.



**Imagen 5.1.1:** Proceso de la metodología del Prototipo Evolutivo.

## 5.2 Definición de prototipos

Se ha contemplado el desarrollo de 4 prototipos que corresponden al número de módulos que integrarán el Clasificador Unimodal de Archivos Musicales.

En la tabla 5.2.1 se enlistan los prototipos con una breve descripción de la función que desempeña cada uno en el sistema.

Prototipo	Función a desempeñar
Prototipo 1: Lectura de archivos WAV de un directorio.	La función de este prototipo será leer todos los archivos en formato WAV que se encuentren en un directorio, y extraer la señal digital del audio musical de cada archivo.
Prototipo 2: Procesamiento de la señal digital de audio.	La señal digital de cada archivo será procesada por este prototipo, se tomará solo un número determinado de muestras y se segmentará la señal en 10 fragmentos o frames de 65,536 muestras cada uno. También será el encargado de la extracción de características del audio. Hará uso de los frames o de la señal completa para el procesamiento y realizar un número establecido de funciones matemáticas y algoritmos para lograr la extracción de las características del audio determinantes para su posterior clasificación.
Prototipo 3: Clasificación de archivos musicales.	De acuerdo al vector de características recibido por el prototipo 2, empleará alguna red neuronal o algún algoritmo de clasificación para lograr clasificar cada archivo musical en alguno de los géneros rock, pop, jazz, electrónica, tropical, clásica y hip hop. Devolverá el género al que fué clasificado cada archivo.
Prototipo 4: Interfaz gráfica, prototipo integrador.	A través de este prototipo se integra la funcionalidad de cada uno de los módulos anteriores para la realizar la clasificación de archivos WAV por género mediante una interfaz gráfica amigable al usuario.

**Tabla 5.2.1:** Prototipos que integrarán el sistema.

### 5.3 Diagrama de Bloques

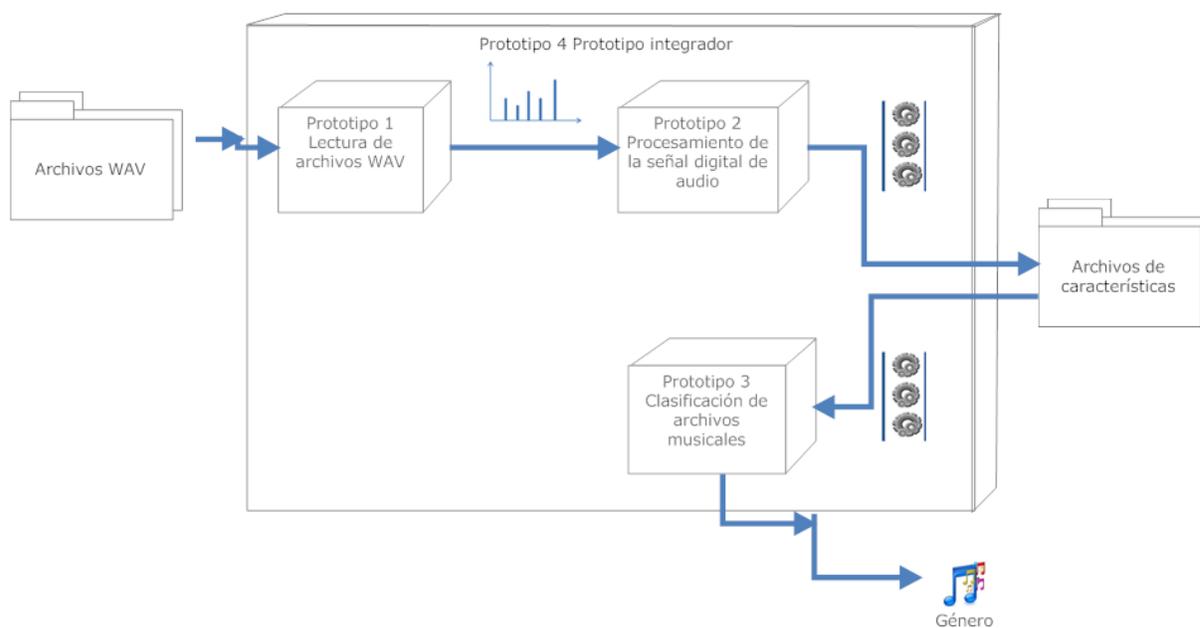
En la Imagen 5.3.1 se muestra de forma general el funcionamiento del sistema. Como entrada al sistema se tiene un conjunto de archivos musicales o un único archivo en formato WAV, que al ser procesados por el Clasificador Unimodal de Archivos son clasificados por género, obteniendo el género al que se clasificó.



**Imagen 5.3.1:** Diagrama de Bloques

## 5.4 Arquitectura general de proyecto

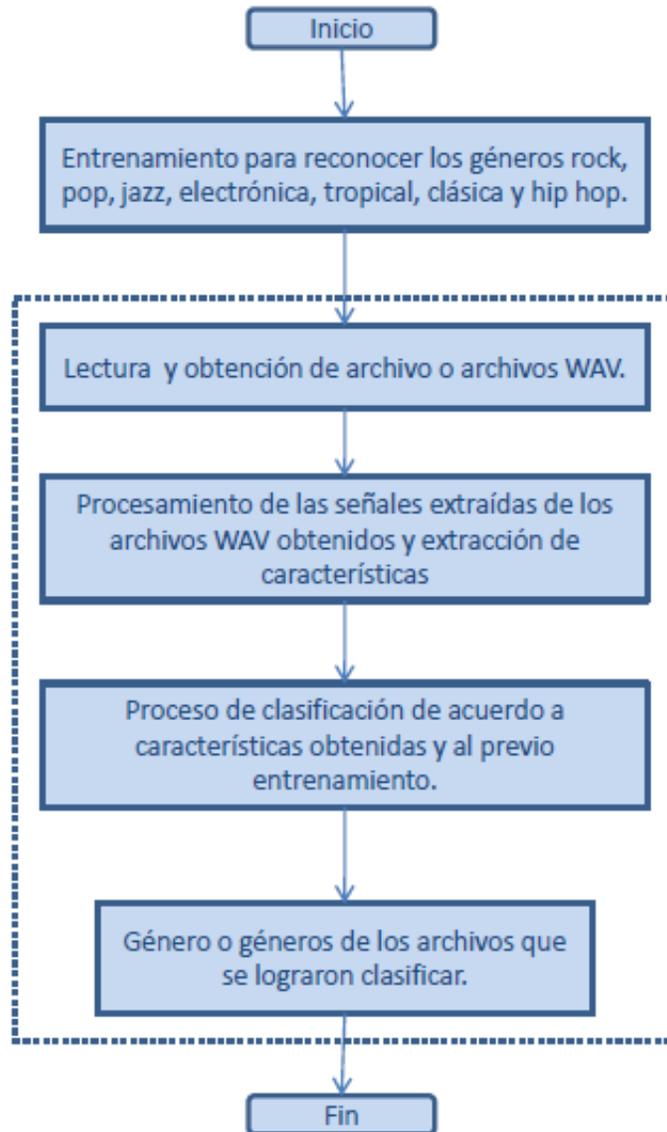
En la siguiente imagen 5.4.1 se muestra la arquitectura general del proyecto.



**Imagen 5.4.1:** Diagrama de arquitectura general.

## 5.5 Diagrama de interacción de módulos

El diagrama 5.5.1 explica la lógica de la interacción general entre los módulos que componen el Clasificador Unimodal de Archivos.



**Imagen 5.5.1:** Diagrama de interacción de módulos

# 6

## Prototipo I: Lectura de archivos WAV en un directorio

El objetivo principal de este prototipo es lograr leer todos los archivos en formato WAV de un directorio y extraer la señal digital de cada archivo de audio, tomando en cuenta que cada archivo WAV del directorio debe ser estandarizado a una velocidad de muestreo de 16,000 muestras por segundo y codificadas con 16 bits.

## 6.1 1º Versión

### 6.1.1 Análisis y diseño

#### Diagrama de bloques

En la imagen 6.1.1 se ilustra el diagrama de bloques del Prototipo 1, en donde se observa que la entrada son los archivos WAV y después de ser procesados por el Prototipo 1 se obtiene un arreglo con las rutas absolutas de los archivos procesados y otro arreglo que contiene la señal digital de los archivos.



**Imagen 6.1.1:** Diagrama a bloques Prototipo 1.

#### Requerimientos funcionales

En la tabla 6.1.1 se enlistan los requerimientos funcionales que se contemplan para el Prototipo 1, Lectura de archivos WAV en un directorio

ID Requerimiento	Requerimiento
ReqF1	El Prototipo 1 leerá todos los archivos cuya extensión sea ".wav" de un directorio.
ReqF2	Se recorrerán uno a uno los archivos del directorio con extensión ".wav" y el formato WAV y se guardará en un arreglo unidimensional la ruta absoluta del mismo.
ReqF3	De cada ruta absoluta obtenida se extraerá la señal digital de audio del archivo WAV.

**Tabla 6.1.1:** Tabla de requerimientos funcionales del Prototipo 1.

## Requerimientos no funcionales

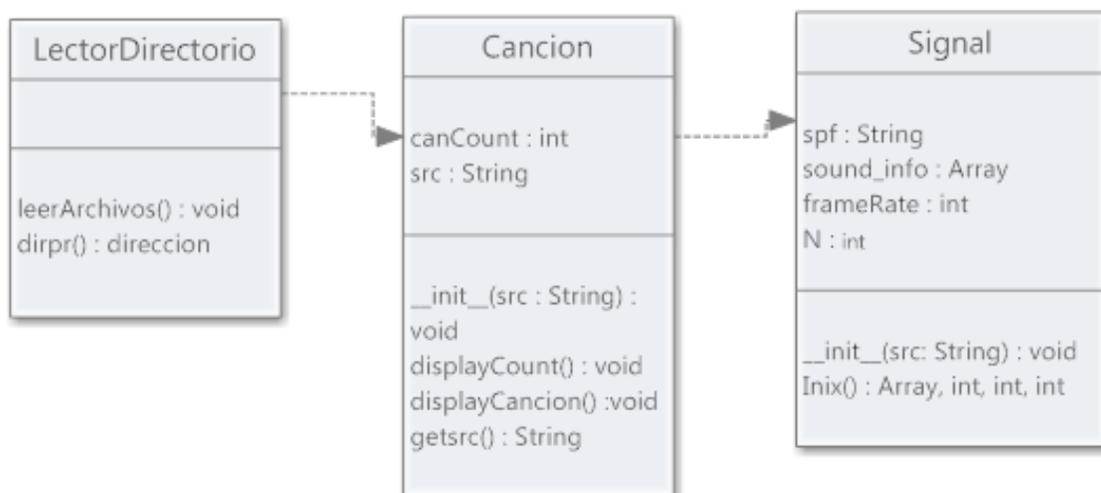
En la tabla 6.1.2 se enlistan los requerimientos no funcionales para el Prototipo 1 Lectura de archivos WAV en un directorio

ReqNoF1	El directorio del que se leerán los archivos .wav será el mismo que en el que se ejecuta el prototipo.
ReqNoF2	La extensión de los archivos de debe ser estrictamente ".wav".
ReqNoF3	La velocidad de muestreo de los archivos debe de ser de 16,000 muestras por segundo, deben de poseer una calidad mono y 16 bits de profundidad.
ReqNoF4	Si no se encuentra ningún archivo con la extensión ".wav", el prototipo enviará un mensaje para indicar que no existen formatos WAV.

**Tabla 6.1.2:** Tabla de requerimientos no funcionales del Prototipo 1.

## Diagrama de clases

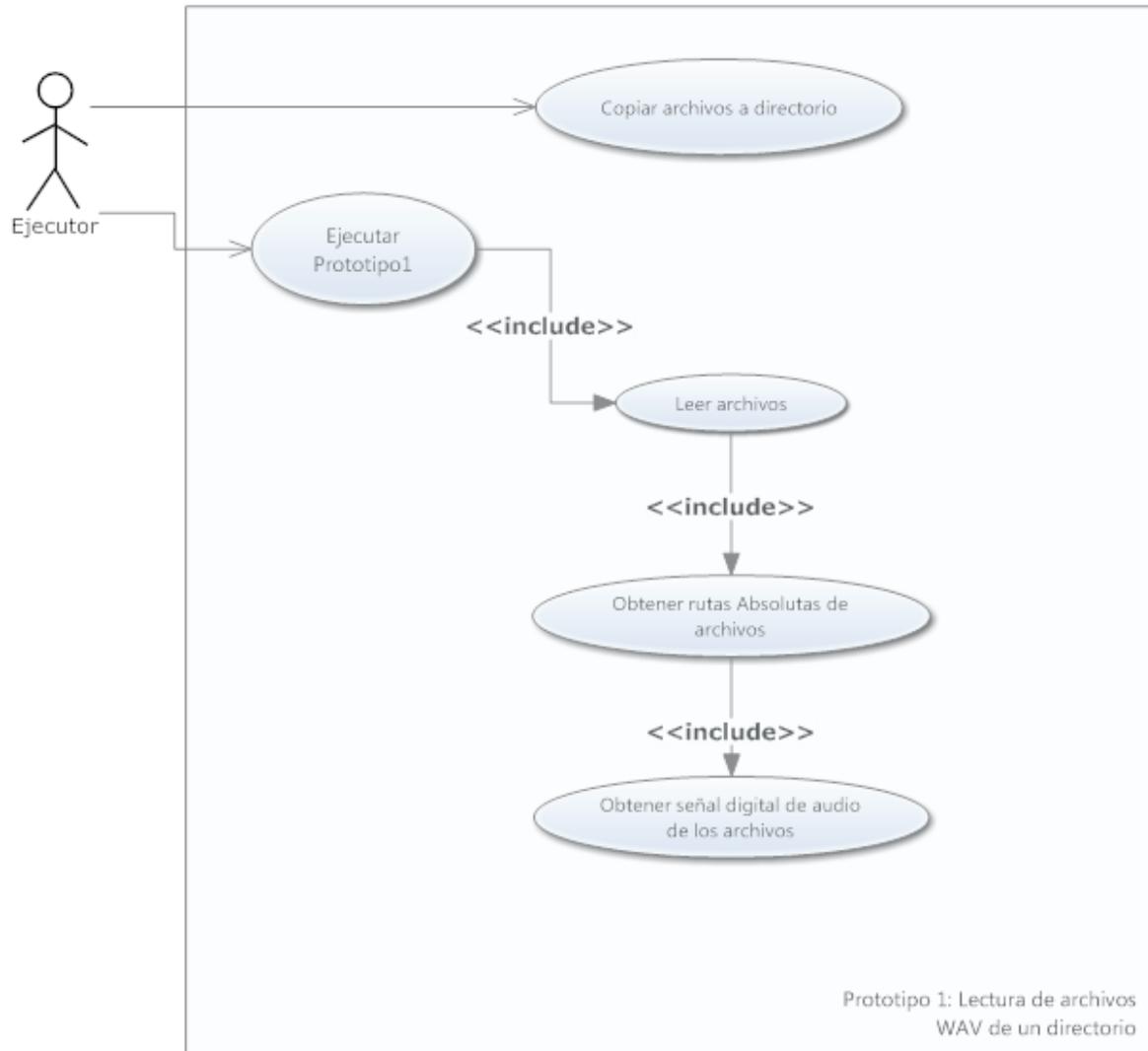
Este prototipo se encuentra conformado por tres clases, LectorDirectorio, Cancion y Signal. El diagrama de clases de la figura 6.1.2



**Imagen 6.1.2:** Diagrama de clases.

## Casos de uso

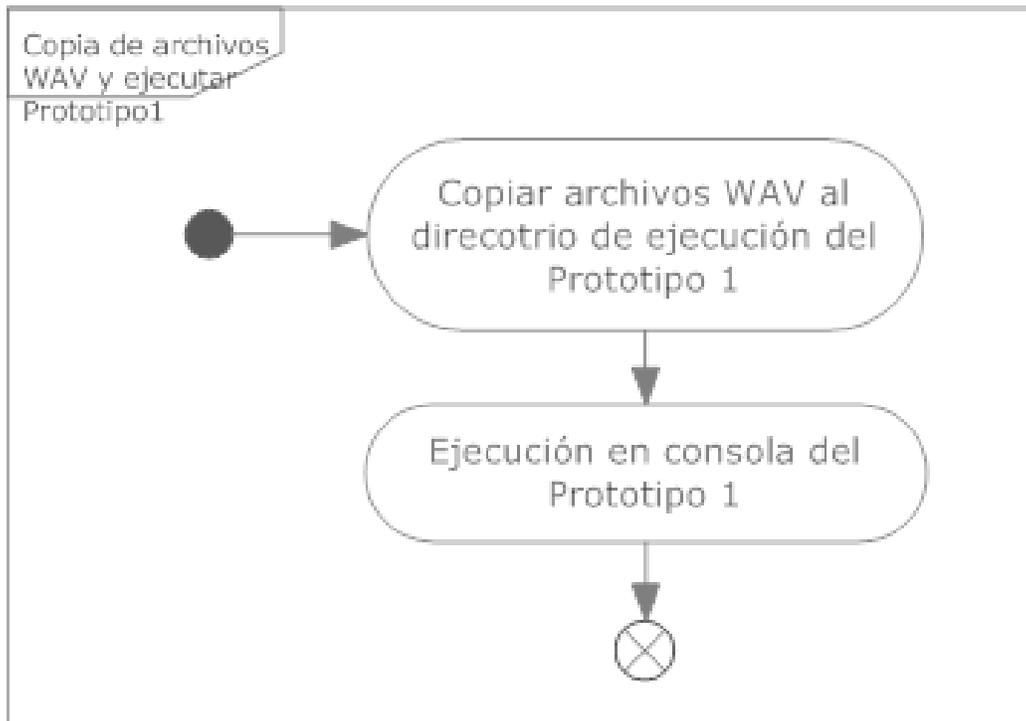
La figura 6.1.3 muestra el diagrama de casos de uso del Prototipo 1: Lectura de archivos WAV en un directorio.



**Imagen 6.1.3:** Diagrama de casos de uso.

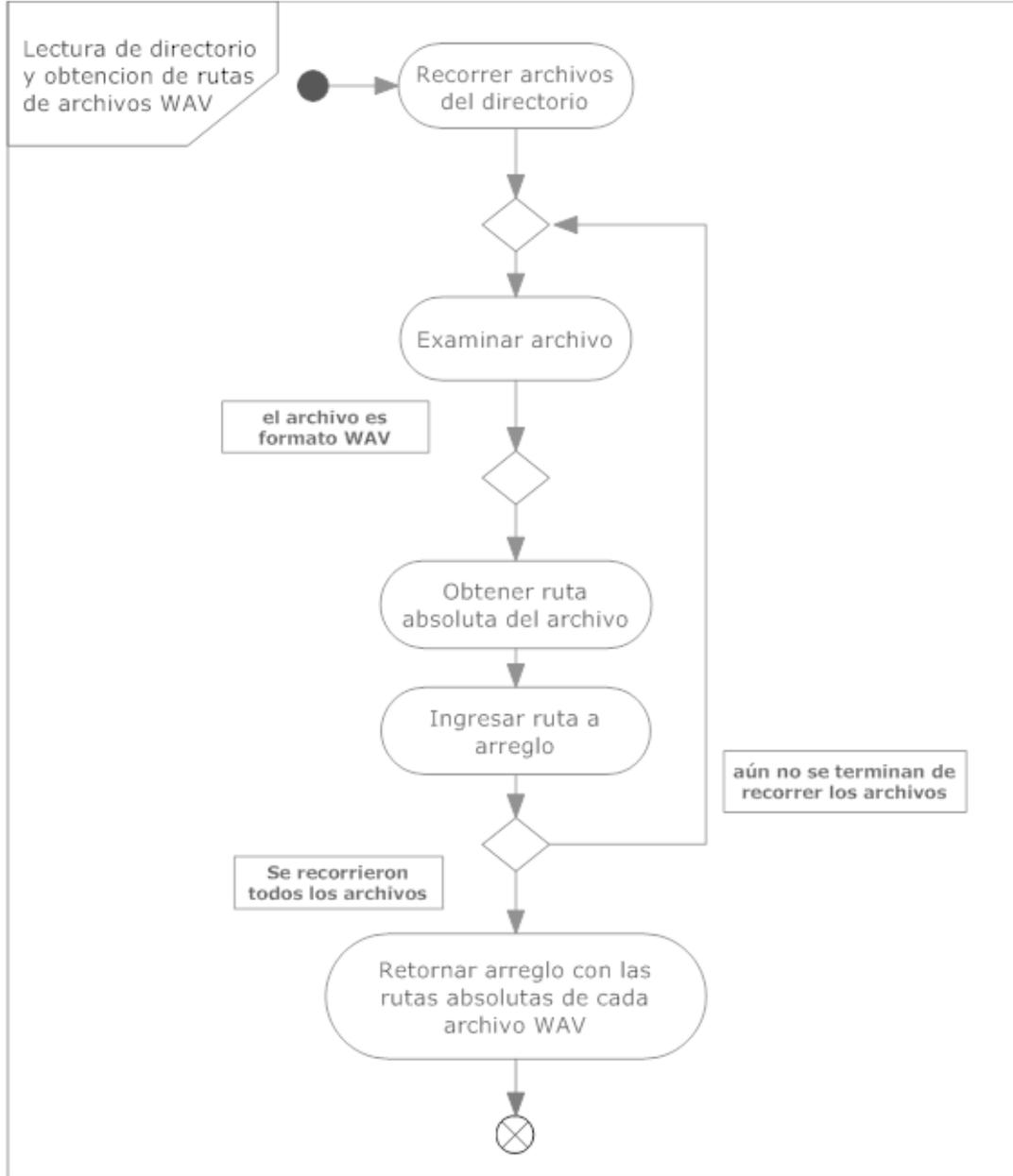
## Diagramas de actividades

En la figura 6.1.4 se muestra el diagrama de actividades correspondiente a los casos de uso Ejecutar Prototipo 1 y Copiar archivos a directorio.



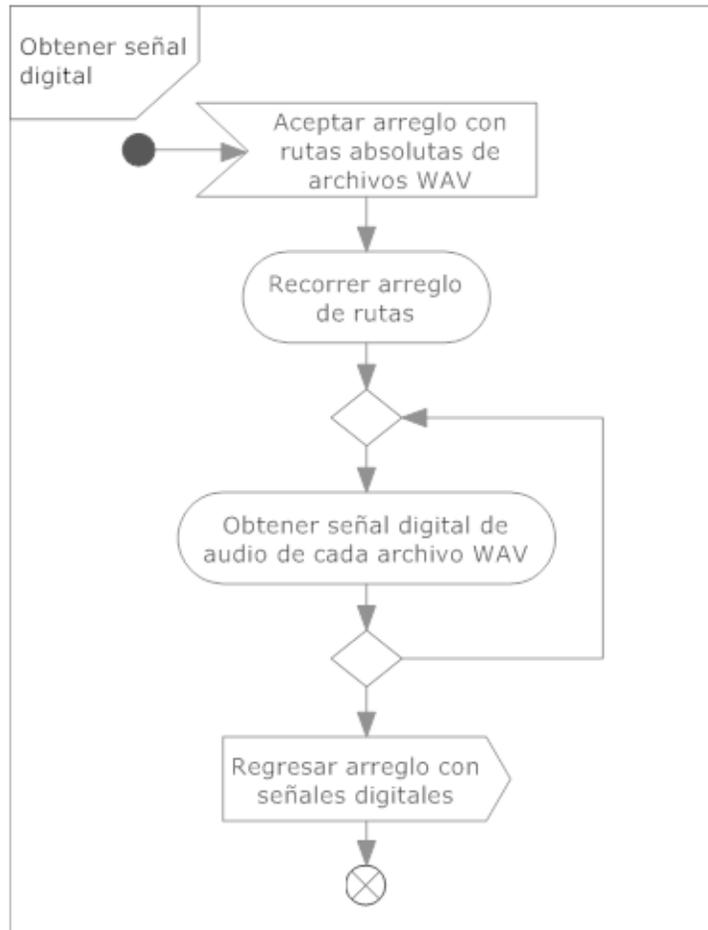
**Imagen 6.1.4:** Diagrama de actividades de ejecución y copia de archivos.

El diagrama 6.1.5 hace referencia a los casos de uso de Leer archivos y Obtener rutas Absolutas de archivos.



**Imagen 6.1.5:** Diagrama de actividades de lectura y obtencion de rutas absolutas de archivos.

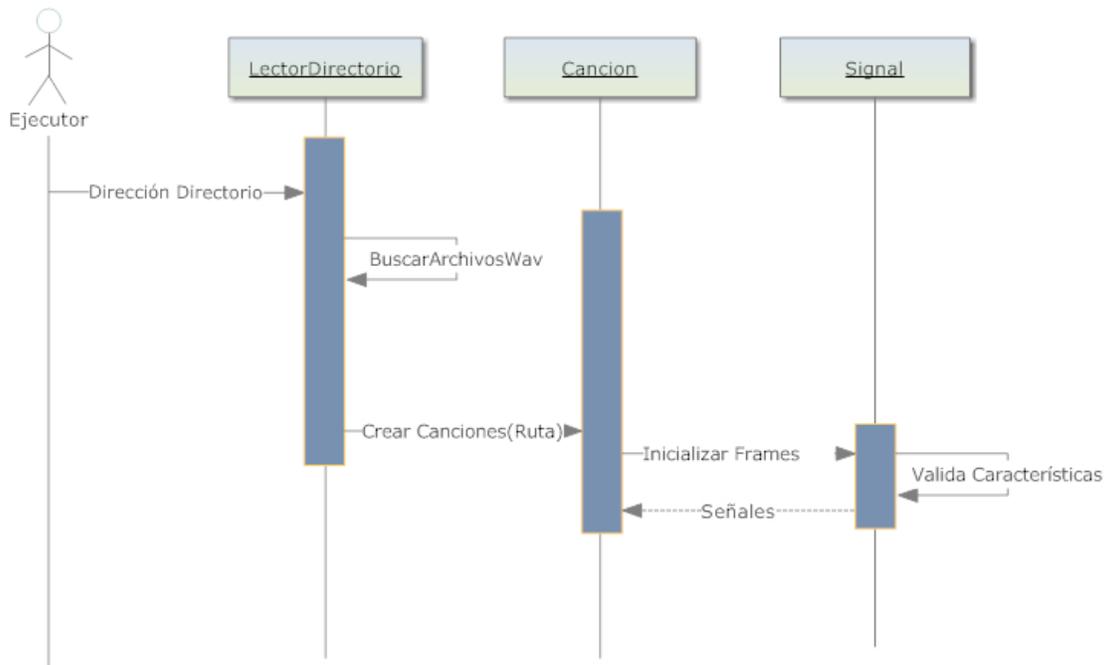
El diagrama 6.1.6 ilustra el caso de uso Obtener señal digital de audio de los archivos.



**Imagen 6.1.6:** Diagrama de actividades de la obtención de la señal digital.

Diagrama de secuencia

En la figura 6.1.7 se ilustra el diagrama de secuencia del Prototipo 1.



**Imagen 6.1.7:** Diagrama de actividades de la obtención de la señal digital.

### 6.1.2 Funcionamiento

- ⇒ Copiar los archivos de audio a la carpeta donde contenga los scripts.
- ⇒ Abrir una terminal y direccionar hacia el directorio que contenga los scripts y las canciones a analizar.
- ⇒ Ejecutar el script principal llamado Principal en consola "python Principal.py".
- ⇒ Se despliega la lista de los archivos wav encontrados en el directorio y se inicializa la señal de cada canción.

En la figura 6.1.8 se ilustra el el resultado del Prototipo 1.

```
mike@ubuntu: ~/Escritorio/Classl
Archivo Editar Ver Terminal Ayuda
mike@ubuntu:~/Escritorio/Classl$ python Principal.py
En la carpeta existen 10 archivos wav
/home/mike/Escritorio/Classl/Strauss - Radetzky March, Op.228 - EMIClassicsUS.wav
/home/mike/Escritorio/Classl/El Amor es Triste.wav
/home/mike/Escritorio/Classl/Beethoven- 'Ode to Joy' from Symphony No. 9 in D minor 'Choral', Op.125 - EMIClassicsUS.wav
/home/mike/Escritorio/Classl/Corazón de Niño.wav
/home/mike/Escritorio/Classl/Depeche Mode - Moonlight Sonata.wav
/home/mike/Escritorio/Classl/Rutter- Cantique De Jean Racine - EMIClassicsUS.wav
/home/mike/Escritorio/Classl/Ballade Pour Adeline - Bang Gur.wav
/home/mike/Escritorio/Classl/Beethoven - Für Elise (from Für Elise - 50 Best Classical Piano) - EMIClassicsUS.wav
/home/mike/Escritorio/Classl/Beethoven - Symphony No. 6 in F_ Shepherds' Hymn - EMIClassicsUS.wav
/home/mike/Escritorio/Classl/Beethoven - Symphony No.5,1st mvt (from Ode to Joy - 50 Best Beethoven) - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/Strauss - Radetzky March, Op.228 - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/El Amor es Triste.wav
Dirección : /home/mike/Escritorio/Classl/Beethoven- 'Ode to Joy' from Symphony No. 9 in D minor 'Choral', Op.125 - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/Corazón de Niño.wav
Dirección : /home/mike/Escritorio/Classl/Depeche Mode - Moonlight Sonata.wav
Dirección : /home/mike/Escritorio/Classl/Rutter- Cantique De Jean Racine - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/Ballade Pour Adeline - Bang Gur.wav
Dirección : /home/mike/Escritorio/Classl/Beethoven - Für Elise (from Für Elise - 50 Best Classical Piano) - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/Beethoven - Symphony No. 6 in F_ Shepherds' Hymn - EMIClassicsUS.wav
Dirección : /home/mike/Escritorio/Classl/Beethoven - Symphony No.5,1st mvt (from Ode to Joy - 50 Best Beethoven) - EMIClassicsUS.wav
```

**Imagen 6.1.8:** Funcionamiento Prototipo 1

# 7

## Prototipo II: Procesamiento de la señal digital de audio

La función que desarrolla este prototipo es la extracción de las características de la señal de audio de los archivos WAV, después estas características se persisten en archivos de texto plano con un formato especial, ya sea el archivo de entrenamiento o un archivo para pruebas.

En la imagen 9 se ilustra el diagrama de bloques del Prototipo II, en el podemos observar que la entrada es un arreglo con las señales de los archivos WAV y después de ser procesados por el Prototipo II se obtienen archivos que contienen las características calculadas, ya sea un archivo para entrenamiento o para pruebas.



**Imagen 9:** Diagrama a bloques del Prototipo II.

- ⇒ Se reciben las señales de audio.
- ⇒ Se calculan sus características.
- ⇒ Se pregunta por el género al que pertenecen las características (para la etapa de entrenamiento).
- ⇒ Se crean y se escriben las características calculadas en archivos de texto plano.
- ⇒ Se genera un archivo de texto único que contiene todas las características de todos los archivos con una respectiva etiqueta que identifica el género (para la etapa de entrenamiento).
- ⇒ Se genera un archivo de texto único que contiene las características de los archivos seleccionados para pruebas (para la etapa de pruebas).

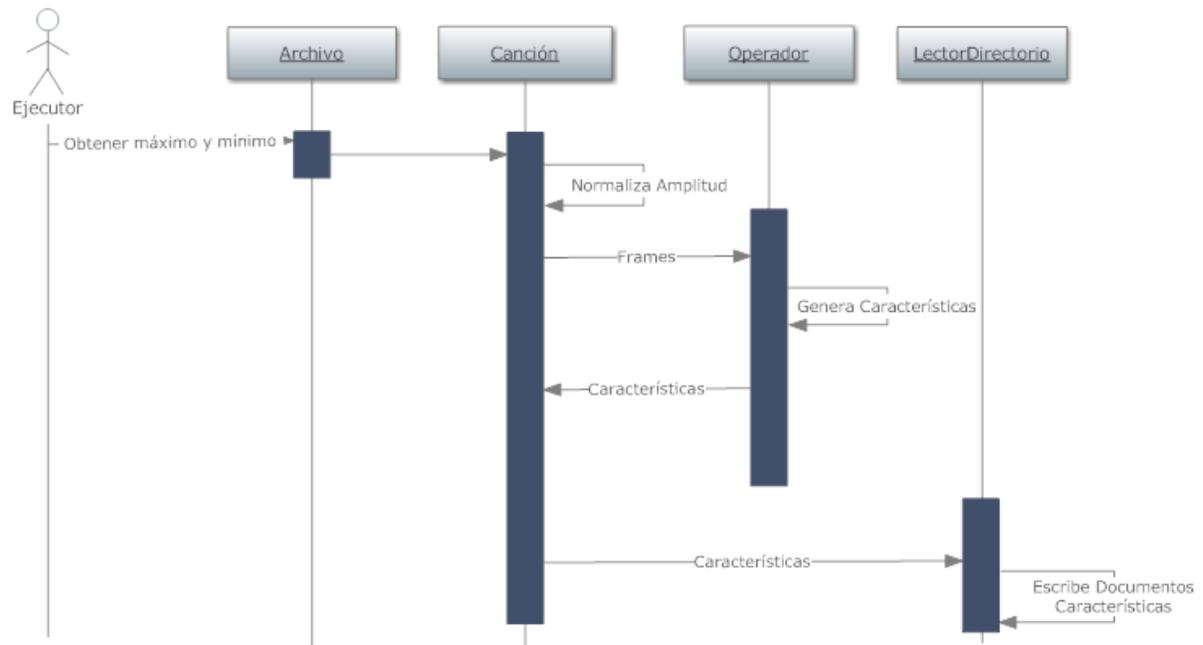
## 7.1 Requerimientos funcionales genéricos

En la tabla 7.1.1 se enlistan los requerimientos funcionales que se contemplan para cada versión del Prototipo II, Procesamiento de la señal digital de audio.

ID Requerimeinto	Requerimiento
ReqF1	El Prototipo II recibe la señal digital extraida previamente por el Prototipo I.
ReqF2	Este prototipo será el encargado de realizar el cálculo de las características que se utilizarán para la clasificación.
ReqF4	Una vez obtenidas las características se persistirán en archivos de texto.

**Tabla 7.1.1:** Tabla de requerimientos funcionales genéricos del Prototipo II.

El siguiente diagrama 7.1.1, corresponde al diagrama de secuencia general del Prototipo II.



**Imagen 7.1.1:** Diagrama de secuencia genérica del Prototipo II.

## 7.2 1º Versión

Esta primera versión del prototipo devuelve un vector de 6 características en el dominio del tiempo. Las características calculadas son volumen, tono, cruce por cero, centro del espectro, flujo y roll off.

### 7.2.1 Requerimientos no funcionales

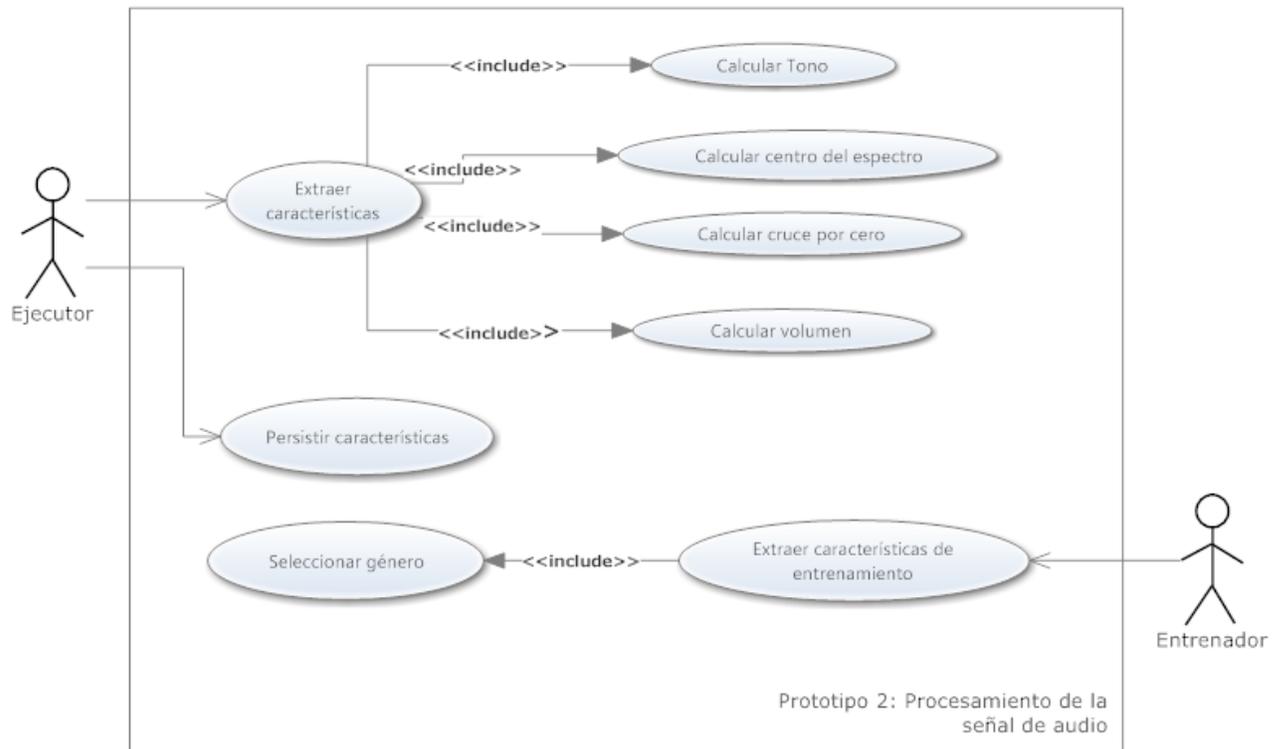
En la tabla 9.2.1 se enlistan los requerimientos no funcionales para la primer versión del Prototipo II.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y cada columna es un elemento de la señal correspondiente a una canción.
ReqNoF2	Para el cálculo de algunas características se contempla la señal completa o en otras la señal se divide en frames o fragmentos. En nuestro caso se toman las primeras 655,360 muestras de la señal de cada archivo. Para el calculo de características en las que se necesita dividir la señal, esta se divide en 10 frames de 65,536 muestras cada uno.
ReqNoF3	Para realizar el cálculo del volumen, tono y cruce por cero, se toma en cuenta toda la señal de 655,360 muestras sin dividir.
ReqNoF4	Para realizar el cálculo del centro del espectro, del flujo y del roll off es necesario dividir la señal de audio digital de 655,360 muestras en 10 ventanas de 65,536 muestras cada una.
ReqNoF5	Se usarán archivos de texto plano para persistir las características calculadas previamente.
ReqNoF6	Para la etapa de entrenamiento se crean 6 archivos de textos, cada uno crresponde a una característica. Estos archivos son guardados en carpetas dependiendo del género al que pertenecen.
ReqNoF7	Para la etapa de pruebas se crea un solo archivo de texto, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta Test.pyData.
ReqNoF8	El archivo correspondiente a la característica del cruce por cero llevará el nombre de CruzexCero.txt.
ReqNoF9	El archivo correspondiente a la característica del tono llevará el nombre de Tono.txt.
ReqNoF10	El archivo correspondiente a la característica del flujo llevará el nombre de Flujo.txt.
ReqNoF11	El archivo correspondiente a la característica del roll off llevará el nombre de RollOff.txt.
ReqNoF12	El archivo correspondiente a la característica del volumen llevará el nombre de Volumen.txt.
ReqNoF13	El archivo correspondiente a la característica del centro del espectro llevará el nombre de CentroEspectro.txt.

**Tabla 7.2.1:** Tabla de requerimientos no funcionales de la versión 1 del Prototipo II.

## 7.2.2 Diagrama de Casos de Uso

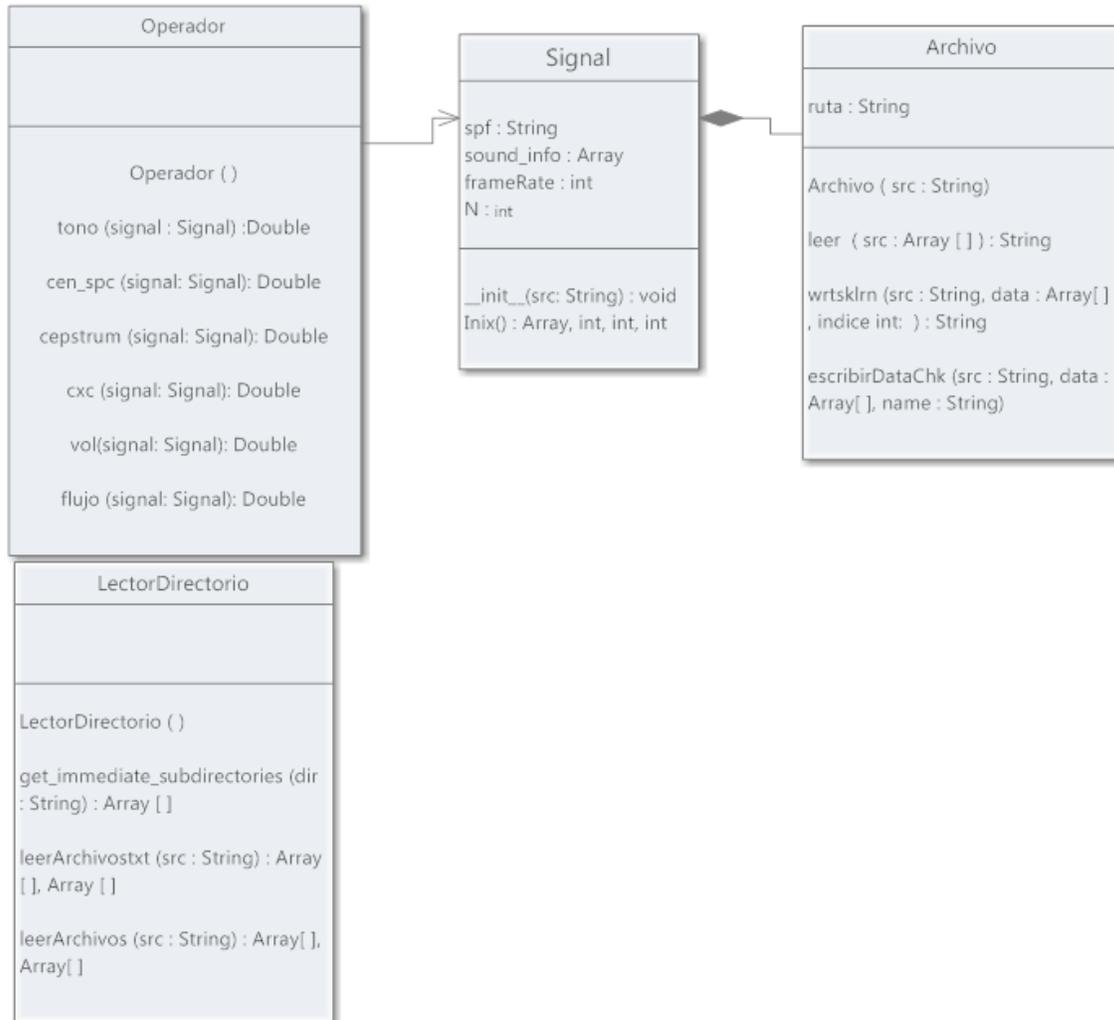
En el diagrama siguiente 7.2.1, se ilustran los casos de uso del Prototipo II versión 1.



**Imagen 7.2.1:** Diagrama de casos de uso del Prototipo II versión 1.

### 7.2.3 Diagrama de Clases

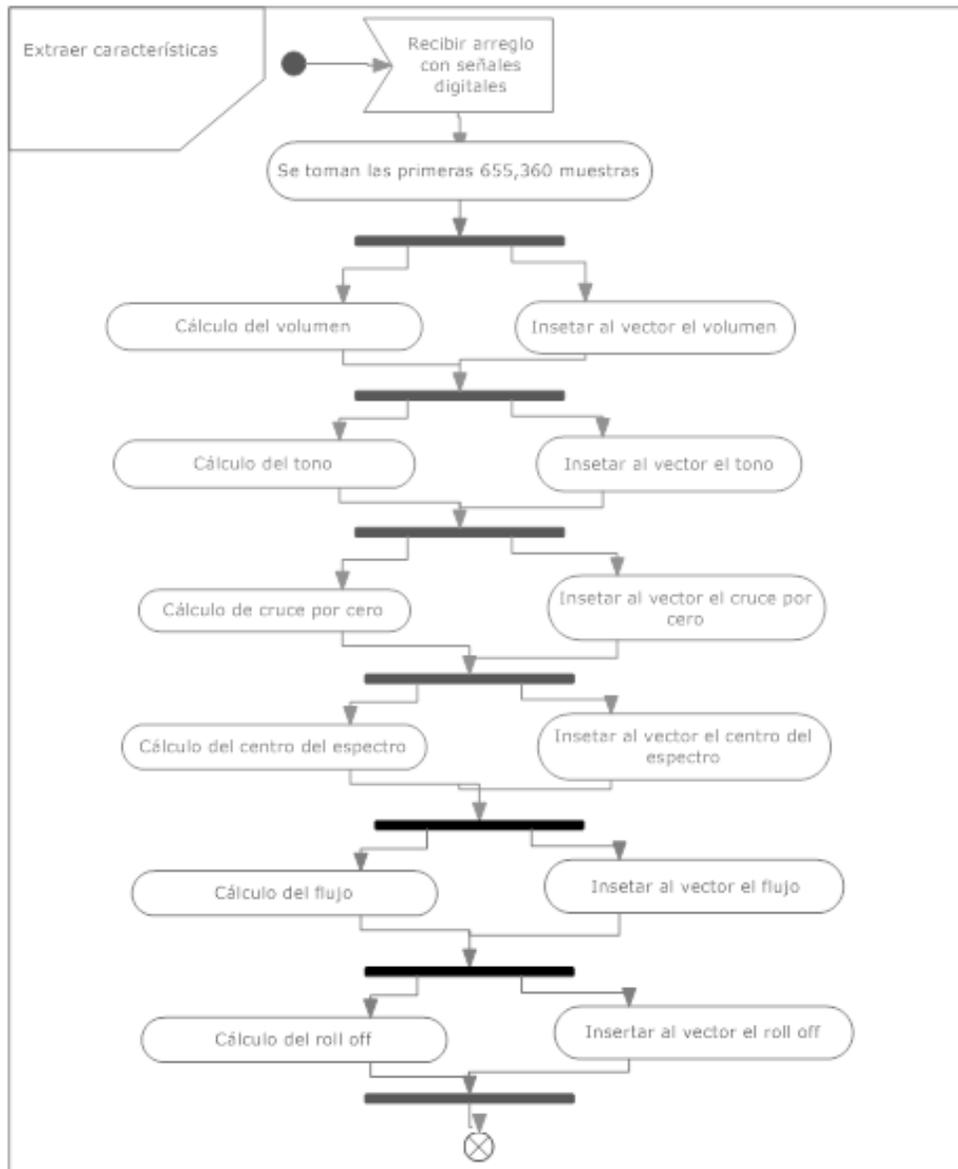
El siguiente diagrama 7.2.2, corresponde al diagrama de clases del Prototipo II versión 1.



**Imagen 7.2.2:** Diagrama de clases del Prototipo II versión 1.

#### 7.2.4 Diagrama de Actividades

El siguiente diagrama 7.2.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 1.



**Imagen 7.2.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 1.

## 7.3 2° Versión

La diferencia entre la versión anterior es que se estandarizan en el dominio del tiempo, de acuerdo a los valores máximos y mínimos de los archivos de entrenamiento los elementos de la señal.

### 7.3.1 Requerimientos no funcionales

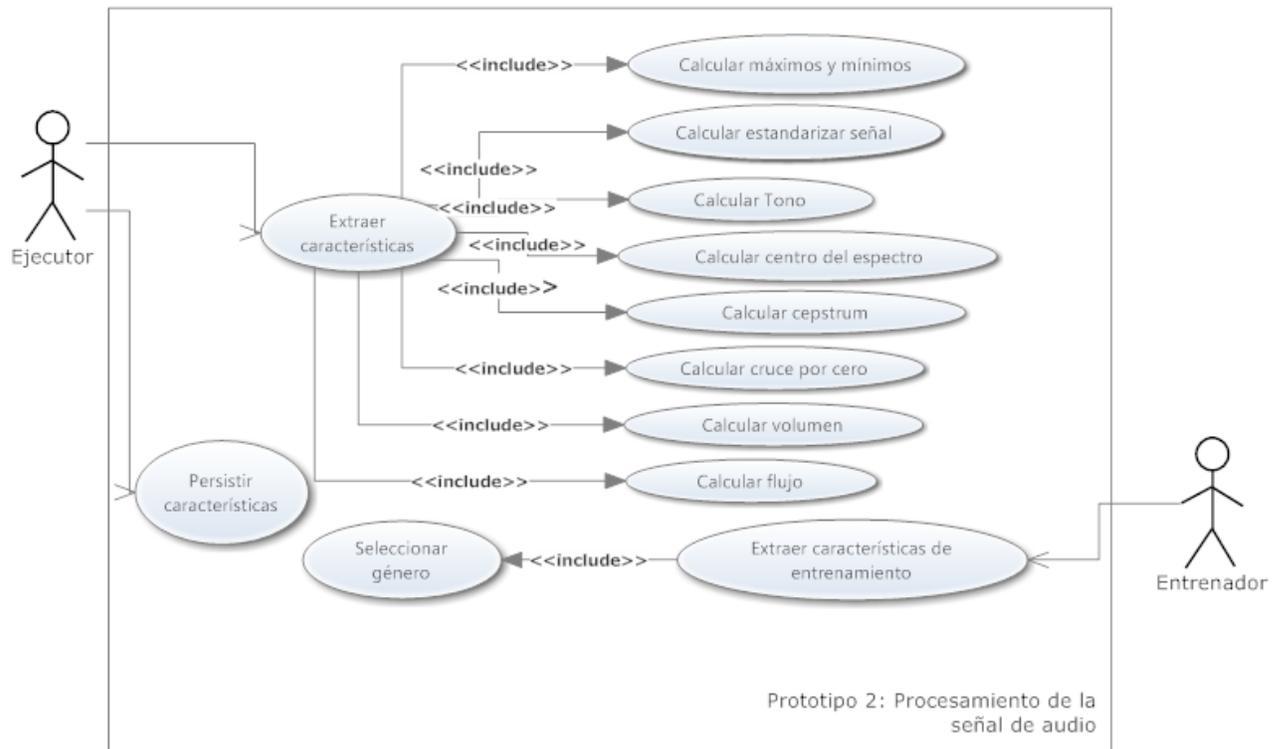
En la tabla 7.3.1 se enlistan los requerimientos no funcionales para la segunda versión del Prototipo II.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de las señales de entrenamiento, ambos valores se almacenan en un archivo de texto plano llamado MaxMin.txt donde cada línea corresponde a un valor.
ReqNoF3	Se utizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados.
ReqNoF4	Para realizar el cálculo del centro del espectro, del flujo es necesario dividir la señal de audio digital de 655,360 muestras en 10 frames de 65,536 muestras cada uno.
ReqNoF5	Para la etapa de pruebas se crea un solo archivo de texto, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta Test.pyData.
ReqNoF6	El archivo correspondiente a la característica del cruce por cero llevará el nombre de CrucexCero.txt.
ReqNoF7	El archivo correspondiente a la característica del tono llevará el nombre de Tono.txt.
ReqNoF8	El archivo correspondiente a la característica del flujo llevará el nombre de Flujo.txt.
ReqNoF9	El archivo correspondiente a la característica del roll off llevará el nombre de RollOff.txt.
ReqNoF10	El archivo correspondiente a la característica del volumen llevará el nombre de Volumen.txt.

**Tabla 7.3.1:** Tabla de requerimientos no funcionales de la segunda versión del Prototipo II.

### 7.3.2 Diagrama de Casos de Uso

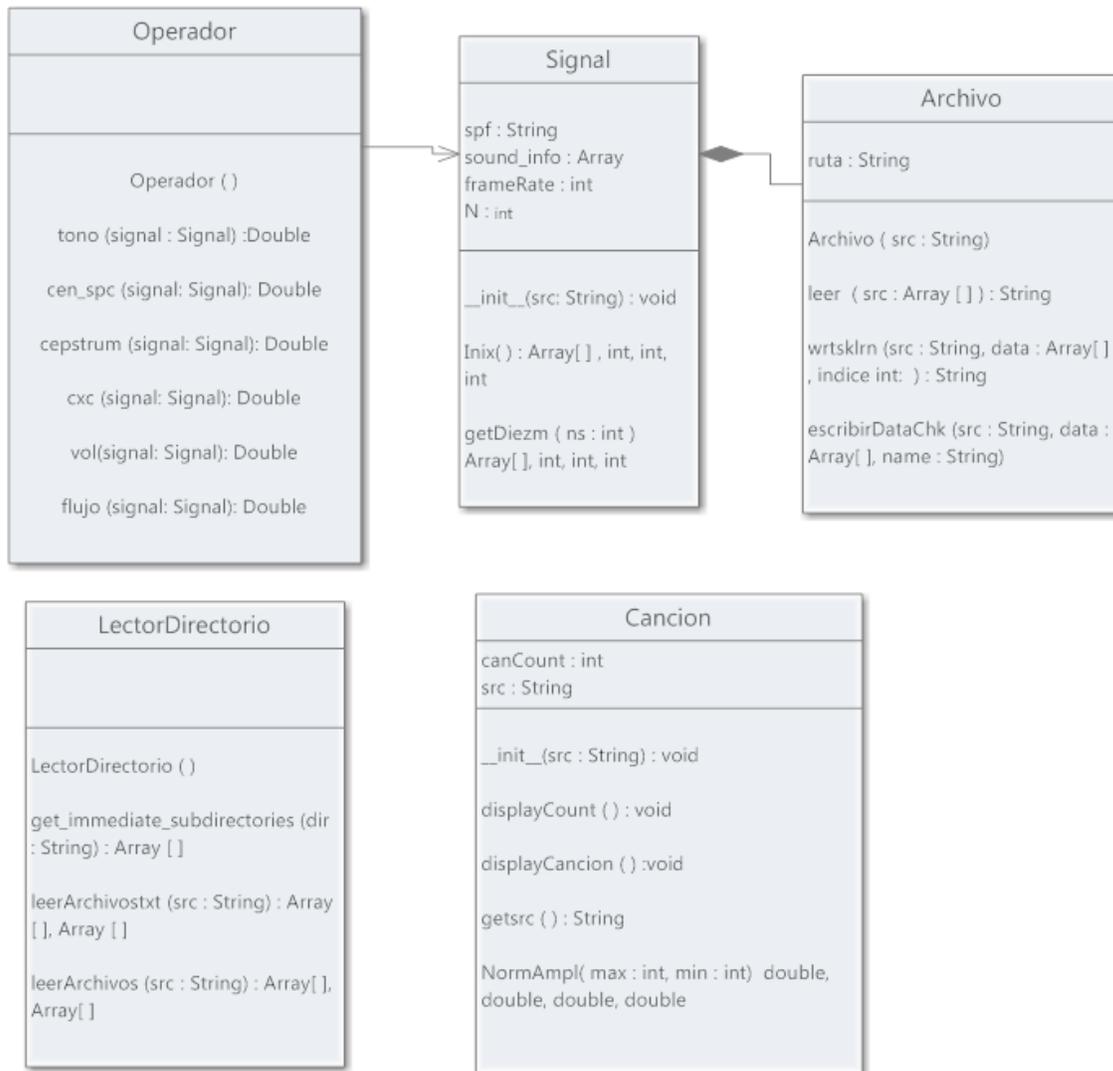
En el diagrama siguiente 7.3.1, se ilustran los casos de uso del Prototipo II versión 2.



**Imagen 7.3.1:** Diagrama de casos de uso del Prototipo II versión 2.

### 7.3.3 Diagrama de Clases

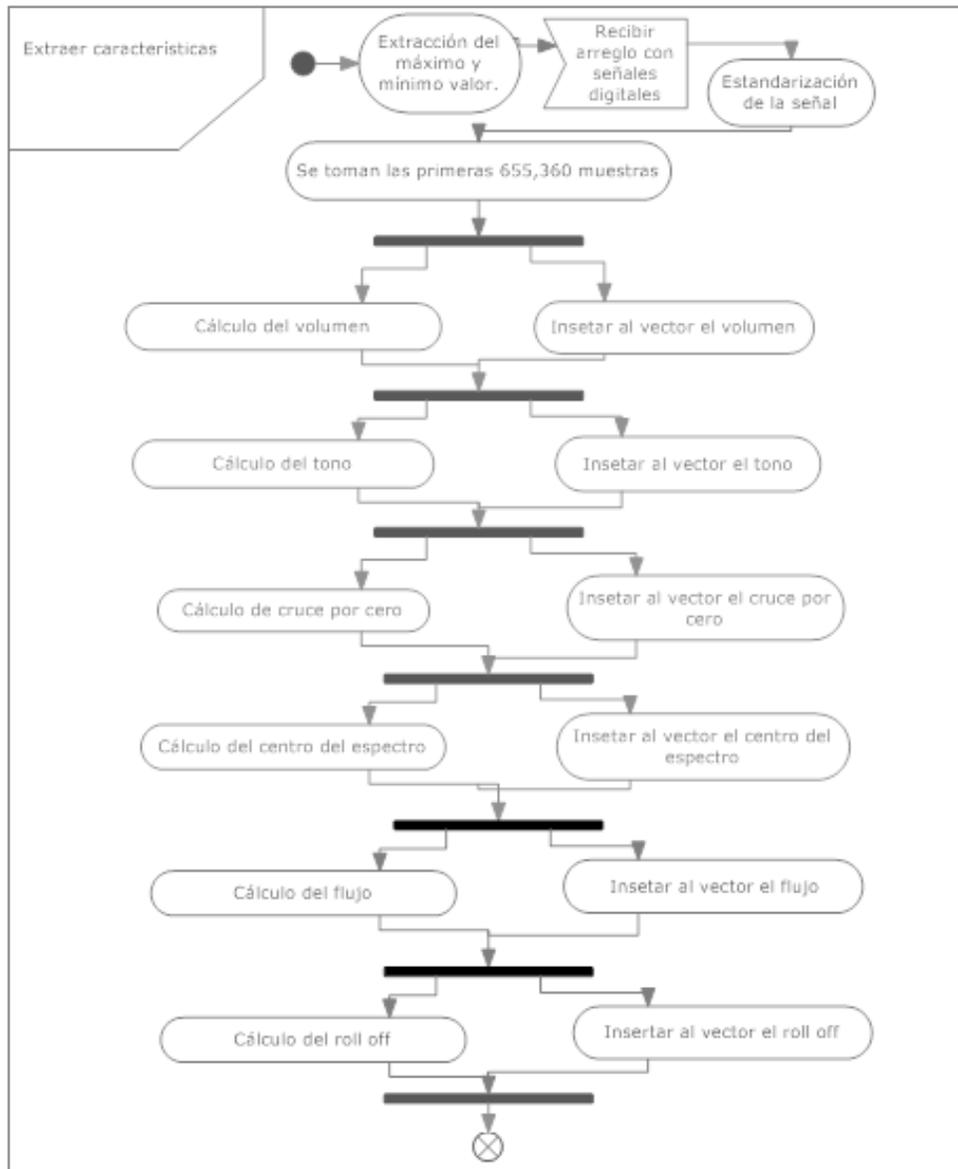
El siguiente diagrama 7.3.2 corresponde al diagrama de clases del Prototipo II versión 2.



**Imagen 7.3.2:** Diagrama de clases del Prototipo II versión 2.

### 7.3.4 Diagrama de Actividades

El siguiente diagrama 7.3.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 2.



**Imagen 7.3.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 2.

## 7.4 3° Versión

Esta versión del prototipo II tiene como finalidad obtener 1024 características, resultado del promedio de la Transformada Rápida de Fourier de ventanas de 1024 muestras cada una.

### 7.4.1 Requerimientos no funcionales

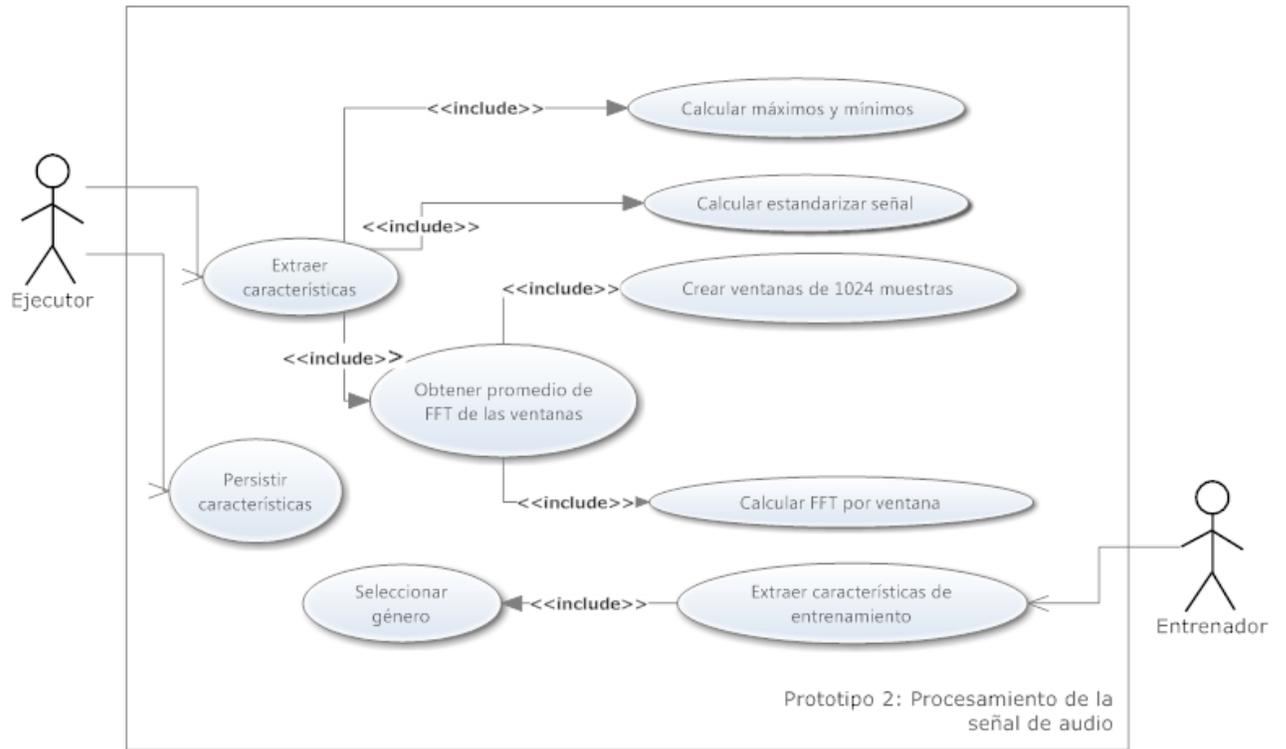
En la tabla 7.4.1, se enlistan los requerimientos no funcionales para el Prototipo II versión 3.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de las señales de entrenamiento, ambos valores se almacenan en un archivo de texto plano llamado MaxMin.txt donde cada línea corresponde a un valor.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Fourier de la señal por ventanas de 1024 muestras. Se hace un promedio del número de muestra $n$ de cada una de las ventanas para formar un vector de 1024 características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF6	El vector de características de cada canción es almacenado en un archivo de formato numpy en un archivo npy.
ReqNoF7	Para la etapa de pruebas se crea un solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta de nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF8	Para la etapa de entrenamiento se crea una carpeta con archivos en formato numpy. Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.

**Tabla 7.4.1:** Tabla de requerimientos no funcionales del Prototipo II versión 3.

#### 7.4.2 Diagrama de Casos de Uso

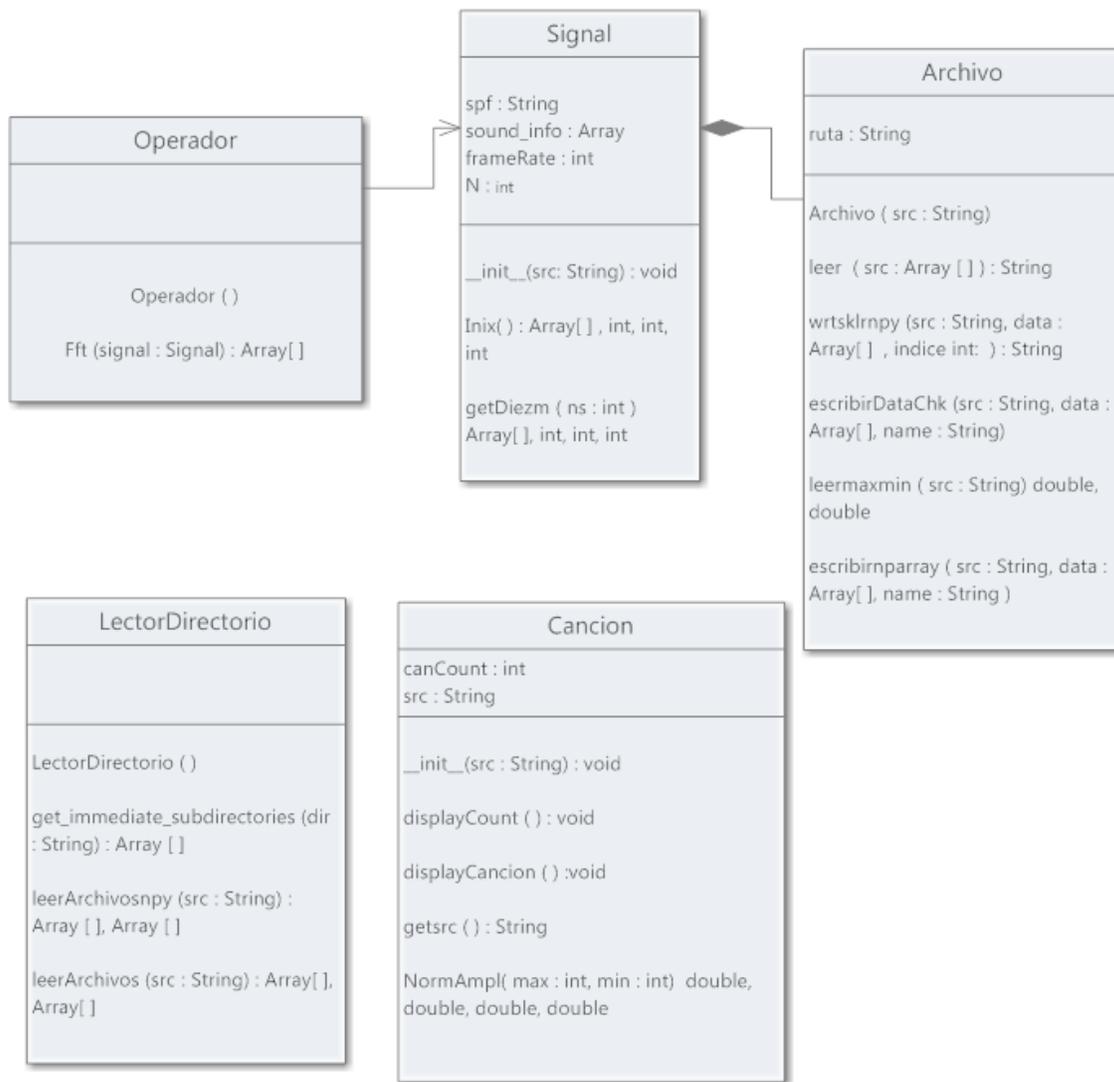
En el diagrama siguiente 7.4.1, se ilustran los casos de uso del Prototipo II versión 3.



**Imagen 7.4.1:** Diagrama de casos de uso del Prototipo II versión 3.

### 7.4.3 Diagrama de Clases

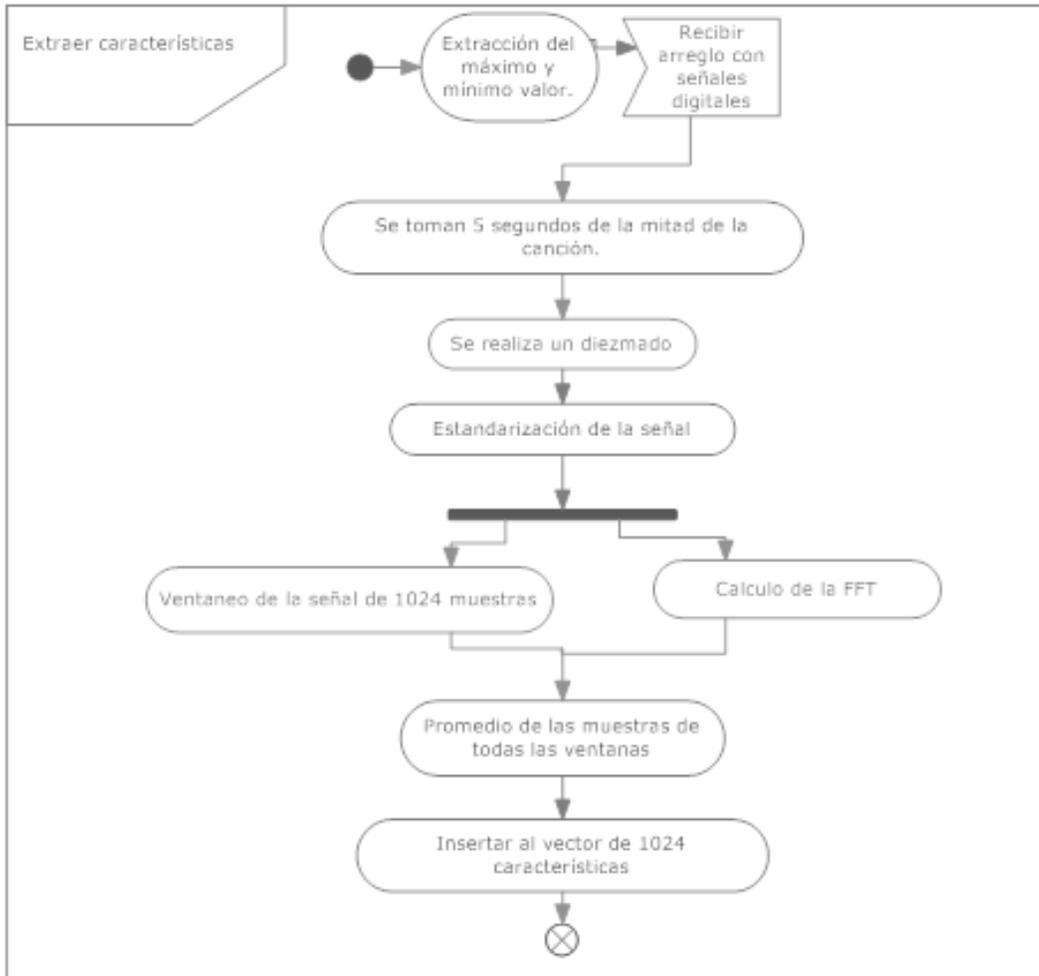
El siguiente diagrama 7.4.2 corresponde al diagrama de clases del Prototipo II versión 3.



**Imagen 7.4.2:** Diagrama de clases del Prototipo II versión 3.

#### 7.4.4 Diagrama de Actividades

El siguiente diagrama 7.4.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 3.



**Imagen 7.4.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 3.

## 7.5 4° Versión

La diferencia entre este prototipo y el anterior es que se forman ventanas de 8 muestras para realizar el cálculo de la FFT, de igual forma se toma la muestra  $n$  de cada ventana y se obtiene un promedio, y así consecutivamente para la muestra  $n$ . El promedio de las 8 muestras conforman el vector de 8 características.

### 7.5.1 Requerimientos no funcionales

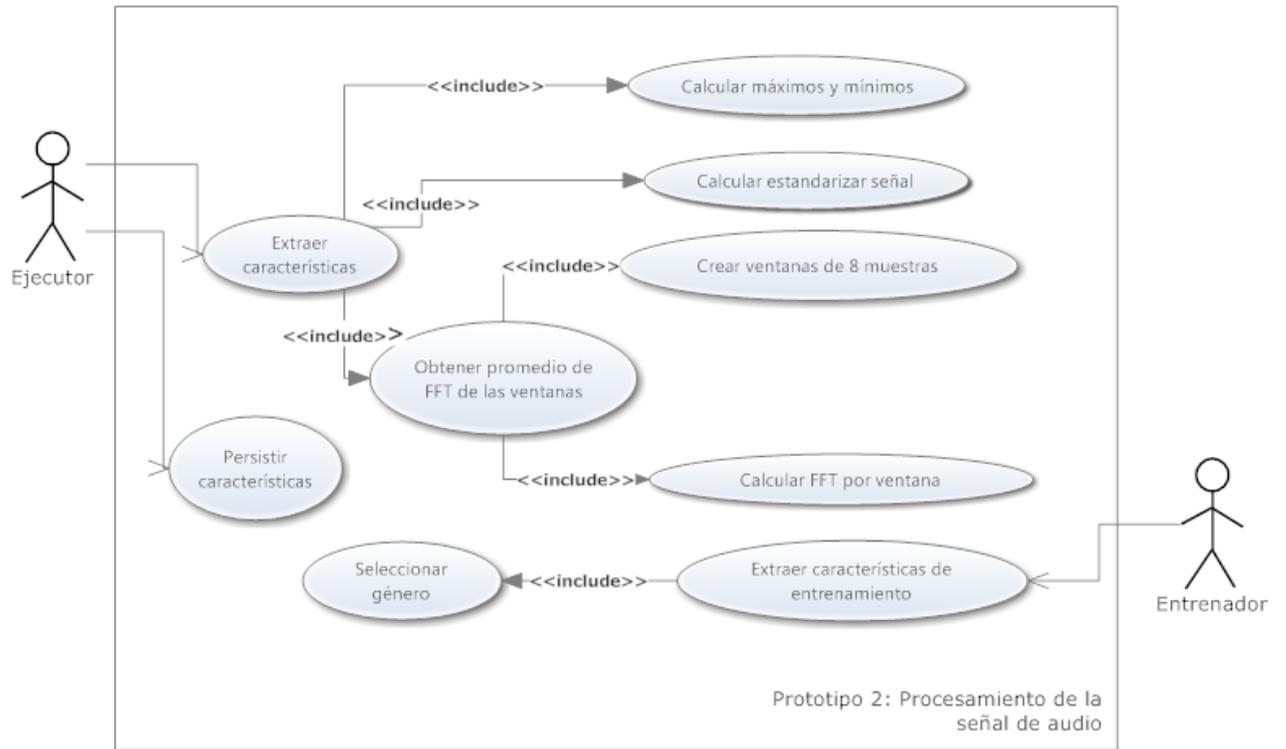
En la tabla 7.5.1 se enlistan los requerimientos no funcionales para el Prototipo II versión 4.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de las señales de entrenamiento, ambos valores se almacenan en un archivo de texto plano llamado MaxMin.txt donde cada línea corresponde a un valor.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Fourier de la señal por ventanas de 8 muestras. Se hace un promedio del número de muestra $n$ de cada una de las ventanas para formar un vector de 8 características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF6	El vector de características de cada canción es almacenado en un archivo de formato numpy en un archivo npy.
ReqNoF7	Para la etapa de pruebas se crea un solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF8	Para la etapa de entrenamiento se crea una carpeta FFTPrincipal.pyData, esta carpeta contiene una carpeta por cada género que se entrena, dentro de la cual están archivos con la extensión npy que contienen las características de cada canción en formato numpy. Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.

**Tabla 7.5.1:** Tabla de requerimientos no funcionales del Prototipo II versión 4.

### 7.5.2 Diagrama de Casos de Uso

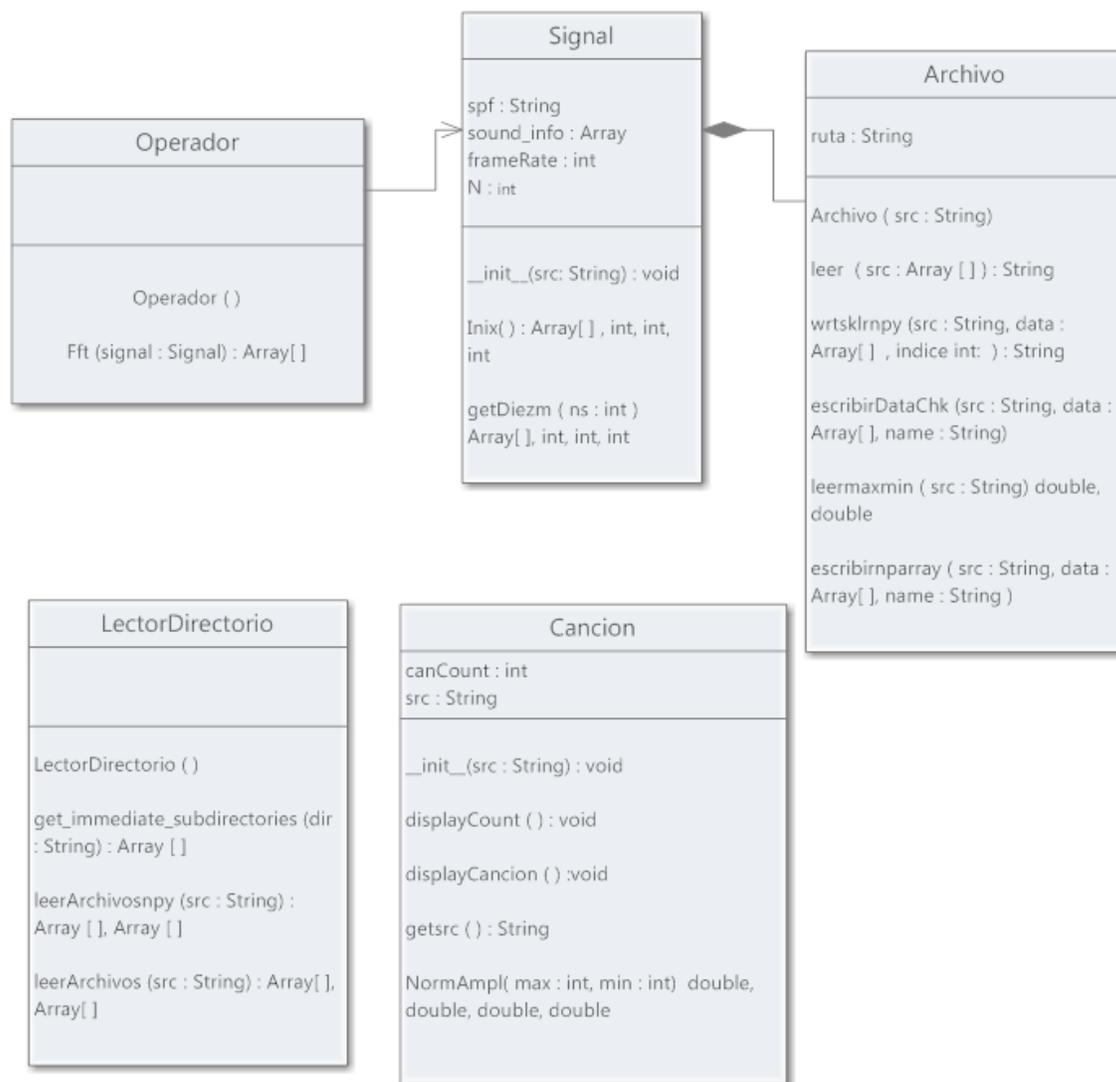
En el diagrama siguiente 7.5.1, se ilustran los casos de uso del Prototipo II versión 4.



**Imagen 7.5.1:** Diagrama de casos de uso del Prototipo II versión 4.

### 7.5.3 Diagrama de Clases

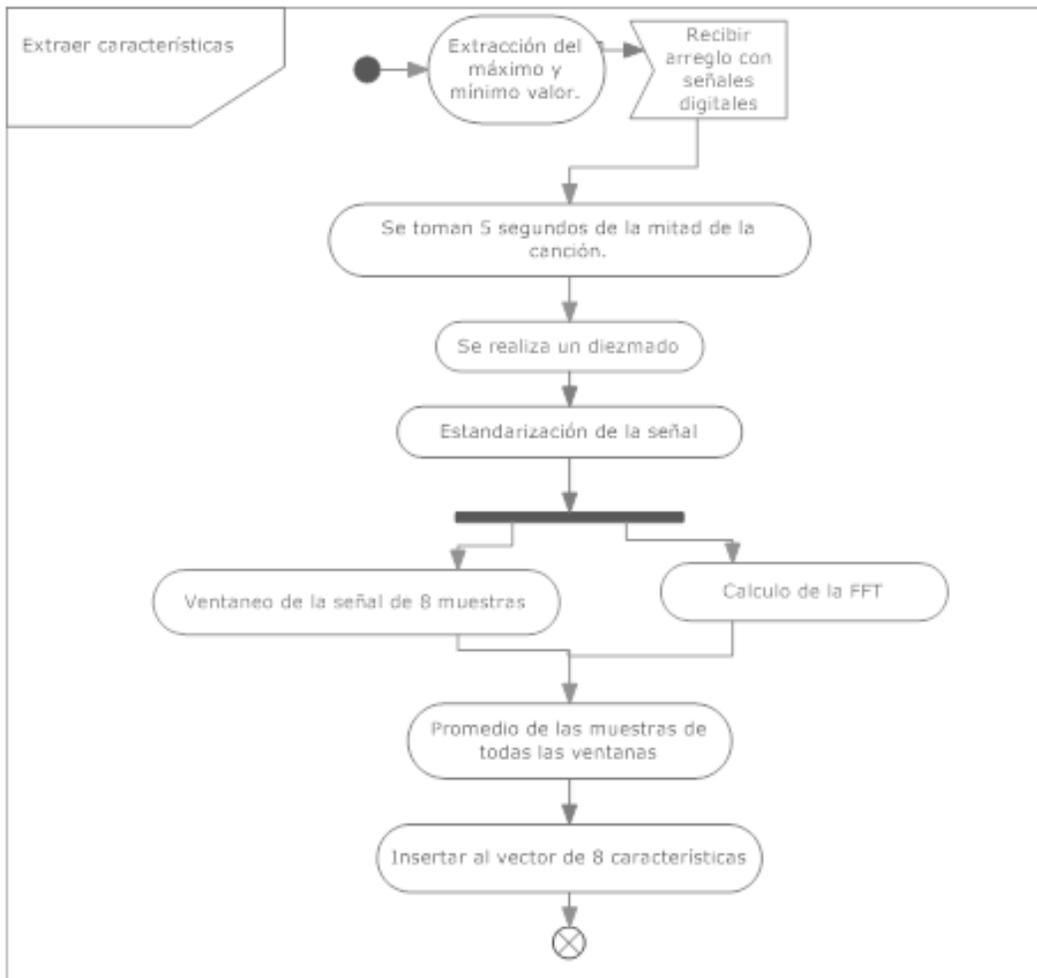
El siguiente diagrama 7.5.2 corresponde al diagrama de clases del Prototipo II versión 4.



**Imagen 7.5.2:** Diagrama de clases del Prototipo II versión 4.

#### 7.5.4 Diagrama de Actividades

El siguiente diagrama 7.5.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 4.



**Imagen 7.5.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 4.

## 7.6 5° Versión

Esta versión arroja un vector de características con dos elementos, ambas características son en el dominio de la frecuencia, la desviación estándar y el promedio, ambas provienen del resultado de la Transformada Rápida de Fourier de la señal de cada canción en el dominio del tiempo.

### 7.6.1 Requerimientos no funcionales

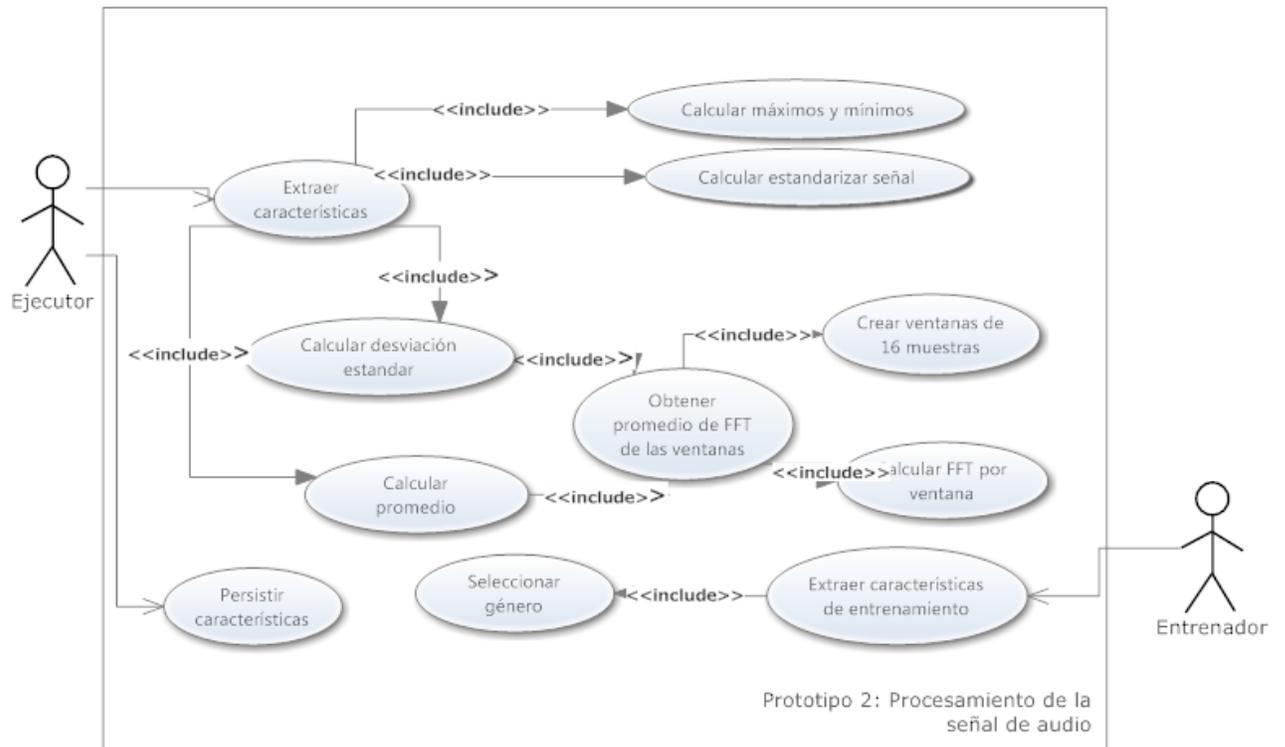
En la tabla 7.6.1 se enlistan los requerimientos no funcionales para el Prototipo II versión 5.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de la señal y se almacena cada valor en una línea de un archivo llamado MaxMin.txt para que posteriormente se puedan estandarizar los valores de la señal de cada archivo musical.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Fourier de la señal por ventanas de 16 muestras. Se hace un promedio del número de muestra $n$ de cada una de las ventanas para formar un vector de 16 características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF6	Se promedia el valor de todos los elementos de la ventana resultante para obtener el primer elemento del vector de características.
ReqNoF7	Se calcula la desviación estándar del valor de todos los elementos de la ventana resultante para obtener el segundo elemento del vector de características.
ReqNoF8	Para la etapa de pruebas se crea 1 solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF9	Para la etapa de entrenamiento se crea 1 carpeta FFTPrincipal.pyData, esta carpeta contiene una carpeta por cada género que se entrena, dentro de la cual están dos archivos de texto plano, DesvStdFft.txt y PromedioFft.txt que contienen las características de la desviación estándar y del promedio respectivamente de cada canción utilizada para el entrenamiento.
ReqNoF10	Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Datatrian.txt en la carpeta FFTPrincipal.pyData.

**Tabla 7.6.1:** Tabla de requerimientos no funcionales del Prototipo II versión 5.

## 7.6.2 Diagrama de Casos de Uso

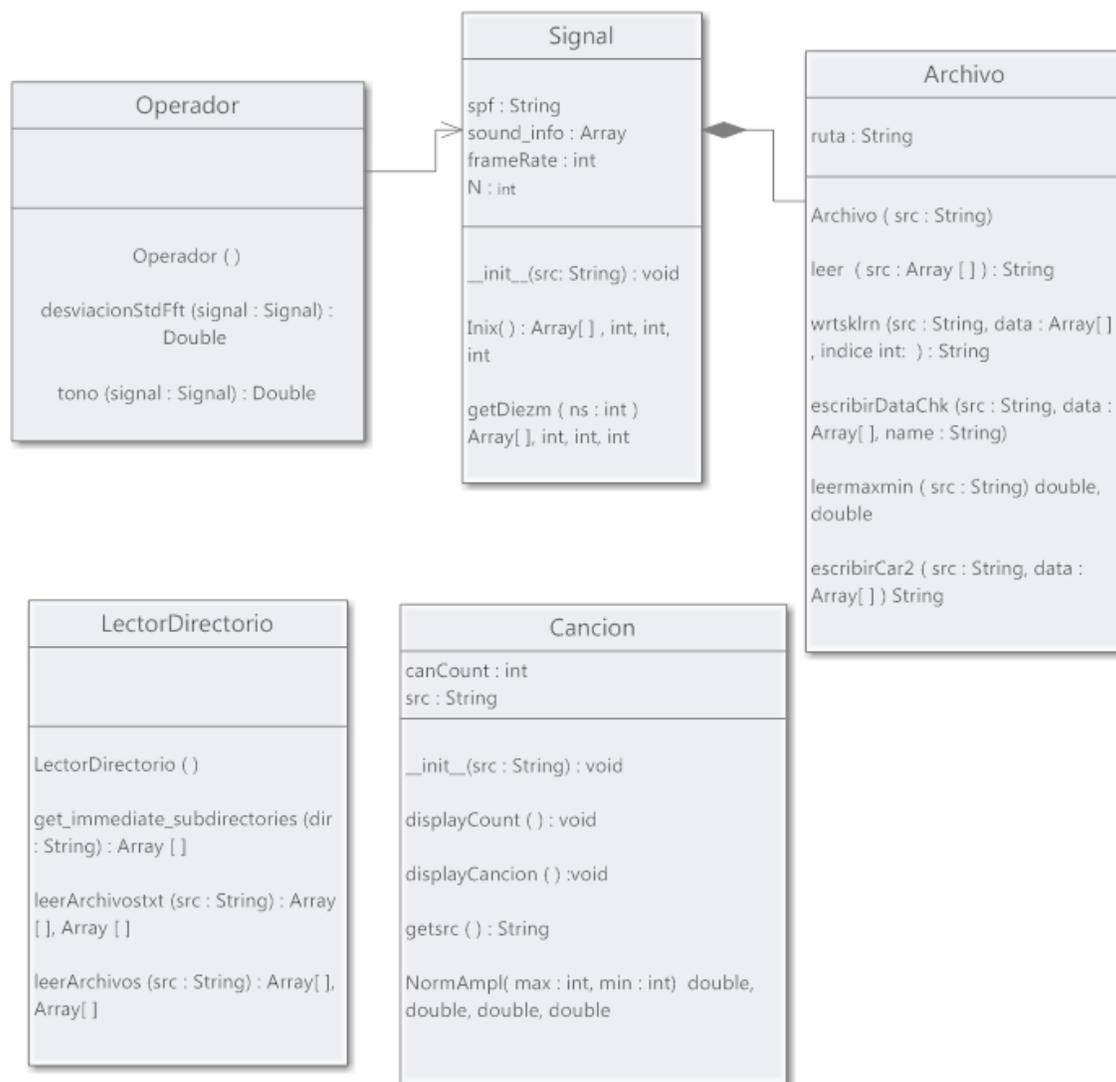
En el diagrama siguiente 7.6.1, se ilustran los casos de uso del Prototipo II versión 5.



**Imagen 7.6.1:** Diagrama de casos de uso del Prototipo II versión 5.

## 7.6.3 Diagrama de Clases

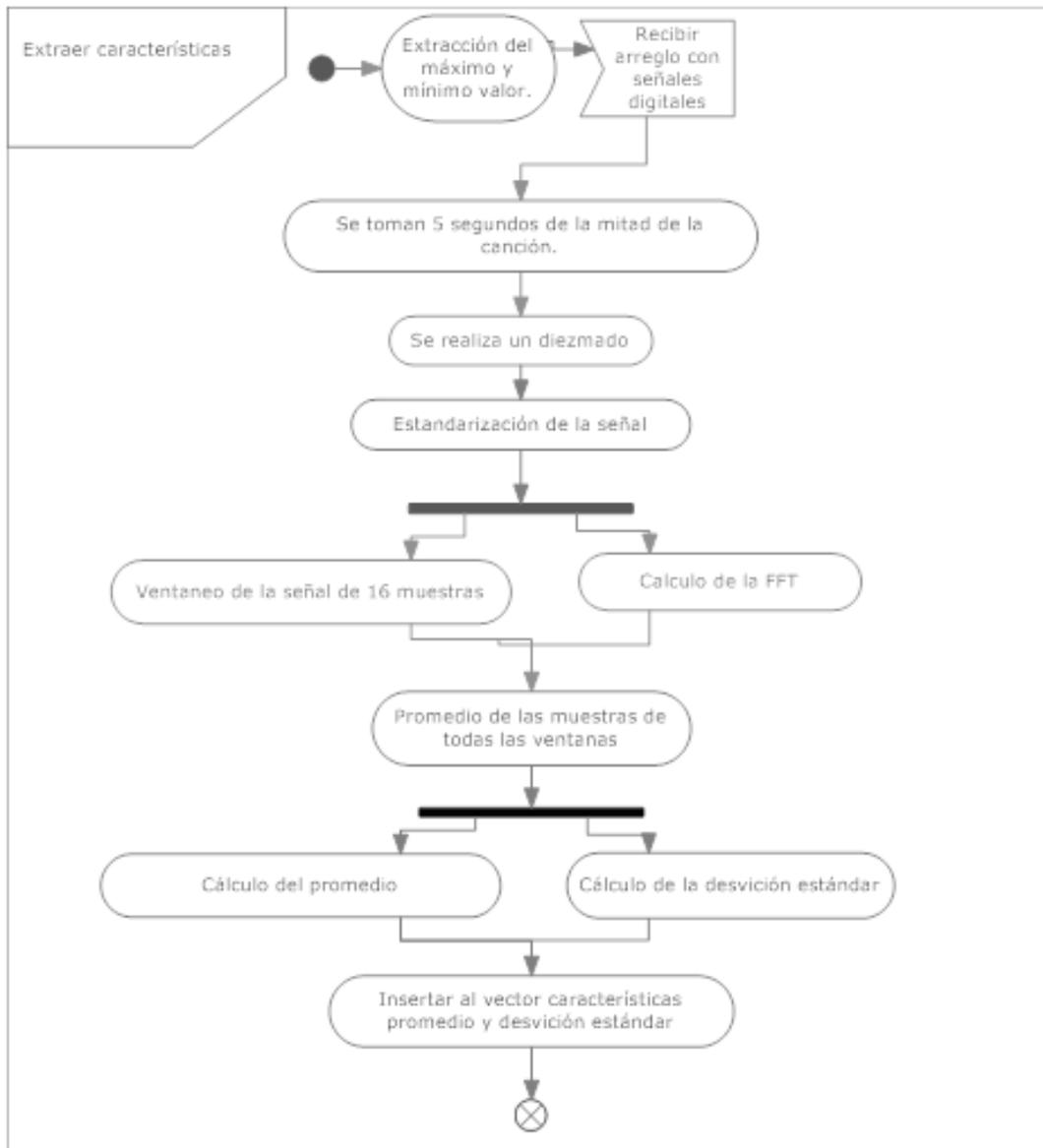
El siguiente diagrama 7.6.2 corresponde al diagrama de clases del Prototipo II versión 5.



**Imagen 7.6.2:** Diagrama de clases del Prototipo II versión 5.

#### 7.6.4 Diagrama de Actividades

El siguiente diagrama 7.6.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 5.



**Imagen 7.6.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 5.

## 7.7 6° Versión

Esta versión del prototipo tiene dos variantes una devuelve un vector de características de 512, 1024 o 2048 elementos. El funcionamiento y el tratamiento que se le da a la señal es muy similar a la versión 4, con la diferencia de que el número de elementos por ventana es el doble del número de elementos del vector de características, es decir, para devolver un vector de 512 elementos se toman ventanas de 1024 muestras, para el vector de 1024 elementos se toman ventanas de 2048 muestras y para el vector de 2048 muestras se utilizan ventanas de 4096 muestras. Esto se debe a que la FFT de la señal es simétrica, de tal forma que resulta redundante y poco eficiente guardar y procesar la otra mitad.

Otro aspecto importante que influye de manera importante en los resultados de la clasificación es que de todas las muestras de entrenamiento se genera un promedio, quedando así una sola muestra, de esta manera la interacción del Prototipo II con el clasificador KNN es muy importante pues el espacio de reconocimiento se ve mucho más limpio pues el número de elementos que existen como referencia es igual al número de géneros con que se entrena.

A continuación en la tabla 7.7.1 se describen los requerimientos no funcionales para esta versión.

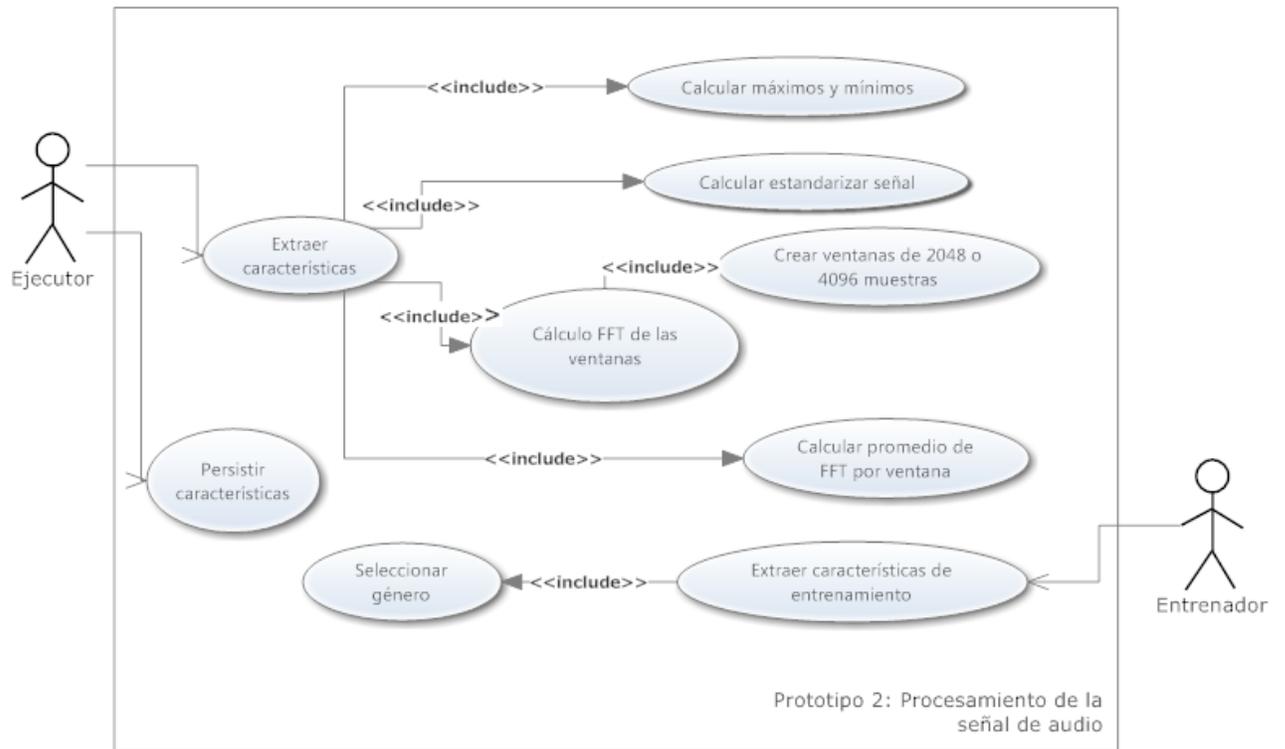
### 7.7.1 Requerimientos no funcionales

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de la señal y se almacena cada valor en una línea de un archivo llamado MaxMin.txt para que posteriormente se puedan estandarizar los valores de la señal de cada archivo musical.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Furier de la señal por ventanas de 1024, 2048 o 4096 muestras.
ReqNoF6	Se toma la mitad de cada ventana, obteniendo así ventanas de 512, 1024 o 2048 muestras ya que la otra mitad es simétrica.
ReqNoF7	Se hace un promedio del número de muestra $n$ de cada una de las ventanas para formar un vector de 1024 o 2048 características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF8	El vector de características de cada canción es almacenada en un archivo de formato numpy en un archivo npy.
ReqNoF9	De todos los vectores de características de las muestras de entrenamiento se promedian cada uno de sus elementos para obtener solo una muestra de referencia sin importar la cantidad de muestras de entrenamiento por género.
ReqNoF10	Para la etapa de pruebas se crea 1 solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF11	Para la etapa de entrenamiento se crea 1 carpeta FFTPrincipal.pyData, esta carpeta contiene una carpeta por cada género que se entrene, dentro de la cual están archivos con la extensión npy que contienen las características de cada canción en formato numpy. Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.

**Tabla 7.7.1:** Tabla de requerimientos no funcionales del Prototipo II versión 6.

### 7.7.2 Diagrama de Casos de Uso

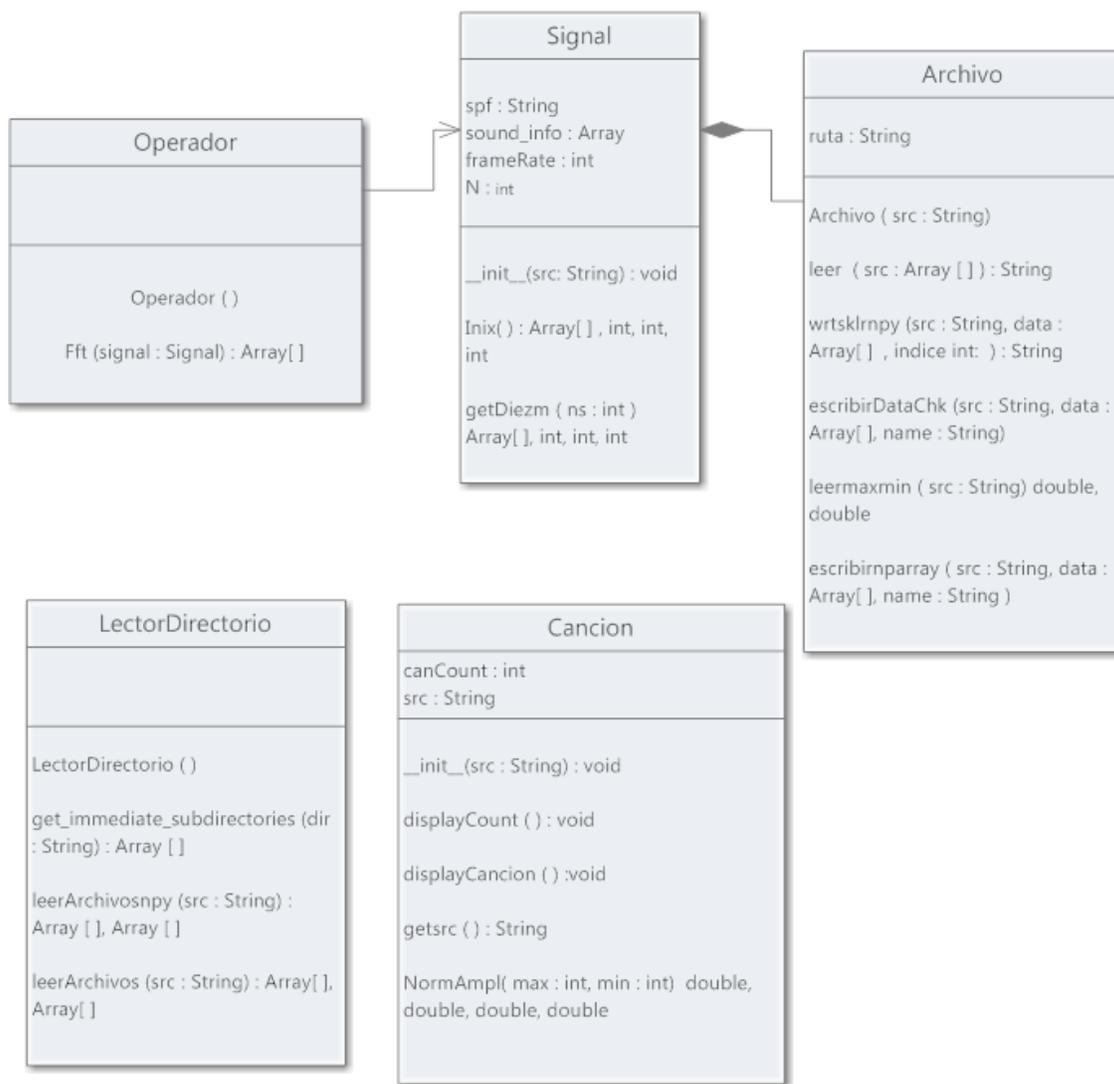
En el diagrama siguiente 7.7.1 se ilustran los casos de uso del Prototipo II versión 6.



**Imagen 7.7.1:** Diagrama de casos de uso del Prototipo II versión 6.

### 7.7.3 Diagrama de Clases

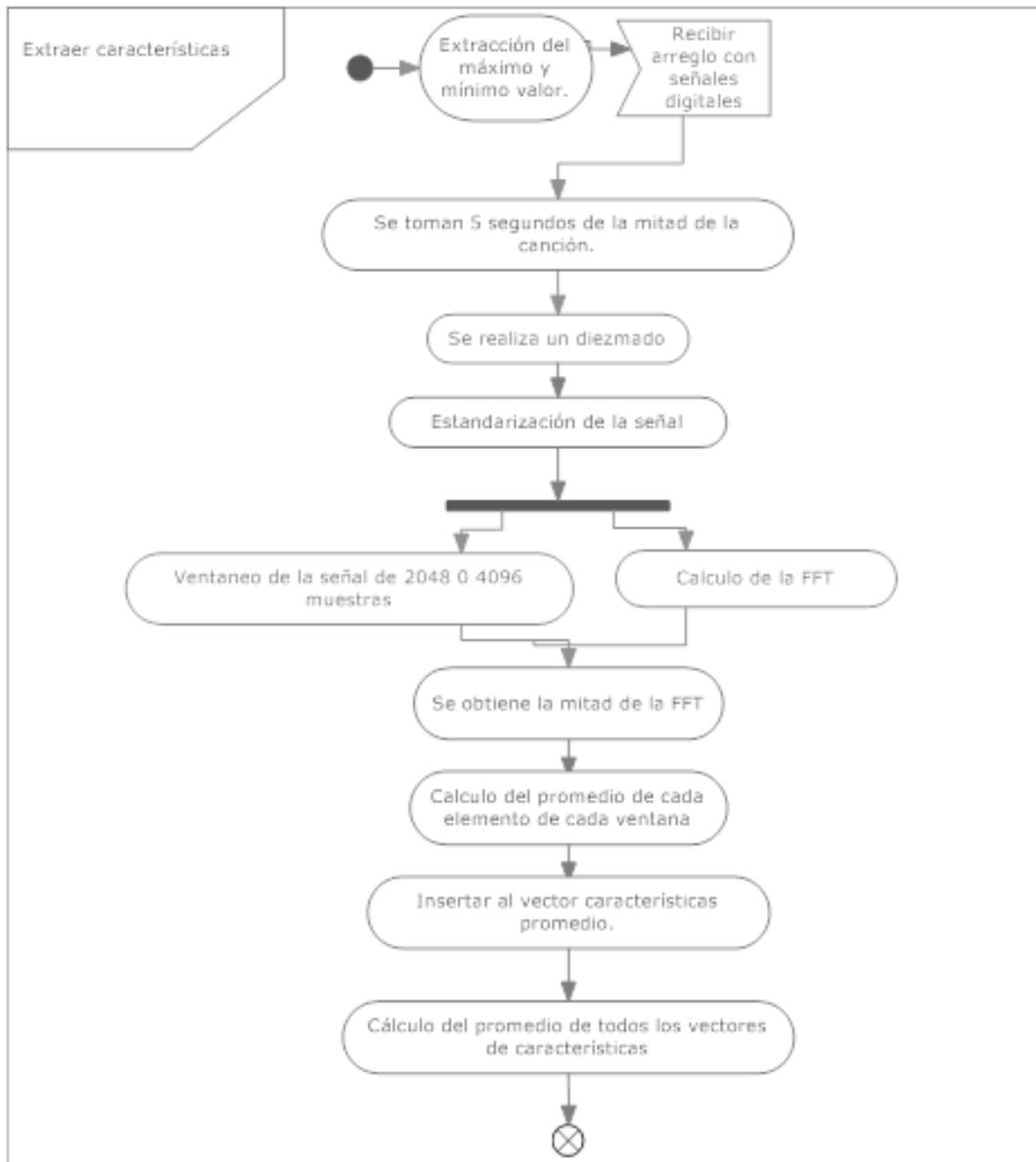
El siguiente diagrama 7.7.2 corresponde al diagrama de clases del Prototipo II versión 6.



**Imagen 7.7.2:** Diagrama de clases del Prototipo II versión 6.

#### 7.7.4 Diagrama de Actividades

El siguiente diagrama 7.7.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 6.



**Imagen 7.7.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 6.

## 7.8 7° Versión

De esta versión existen dos variantes, la primera regresa un vector de características de 8 elementos y la segunda de 16 elementos.

El procesamiento de la señal es similar a la versión 6, pero al final esta versión vuelve a generar un promedio del vector de características de todas las canciones por género.

Otro aspecto importante que influye de manera importante en los resultados de la clasificación es que de todas las muestras de entrenamiento se genera un promedio, quedando así una sola muestra, de esta manera la interacción del Prototipo II con el clasificador KNN es muy importante pues el espacio de reconocimiento se ve mucho más limpio pues el número de elementos que existen como referencia es igual al número de géneros con que se entrena.

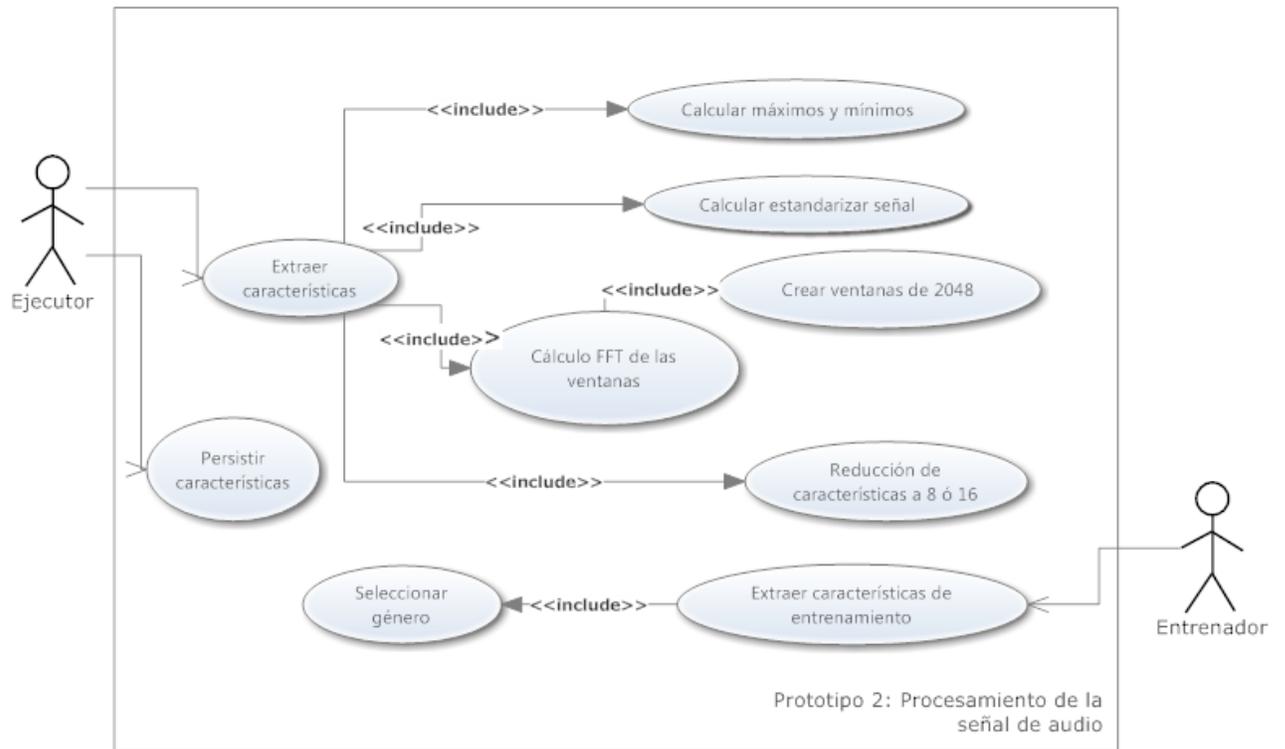
En la tabla 7.8.1 se enlistan los requerimientos no funcionales para el Prototipo II versión 7.

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de la señal y se almacena cada valor en una línea de un archivo llamado MaxMin.txt para que posteriormente se puedan estandarizar los valores de la señal de cada archivo musical.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Furier de la señal por ventanas de 2048 muestras.
ReqNoF6	Se toma la mitad de cada ventana, obteniendo así ventanas de 1024 muestras ya que la otra mitad es simétrica.
ReqNoF7	Se hace un promedio del número de muestra $n$ de cada una de las ventanas para formar un vector de 1024 características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF8	Se forman grupos de 128 o 64 elementos (para el vector de 8 o 16 características). De cada grupo se eliminan los dos puntos con valor más alto y los dos con valor más bajo. El resultado de ese promedio por cada segmento es el vector de características de 8 o 16 elementos.
ReqNoF9	El vector de características de cada canción es almacenada en un archivo de formato numpy en un archivo npy.
ReqNoF10	De todos los vectores de características de las muestras de entrenamiento se promedian cada uno de sus elementos para obtener solo una muestra de referencia sin importar la cantidad de muestras de entrenamiento por género.
ReqNoF11	Para la etapa de pruebas se crea 1 solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta con el nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF12	Para la etapa de entrenamiento se crea 1 carpeta FFTPrincipal.pyData, esta carpeta contiene una carpeta por cada género que se entrene, dentro de la cual estan archivos con la extensión npy que contienen las características de cada canción en formato numpy. Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.

**Tabla 7.8.1:** Tabla de requerimientos no funcionales del Prototipo II versión 7.

### 7.8.1 Diagrama de Casos de Uso

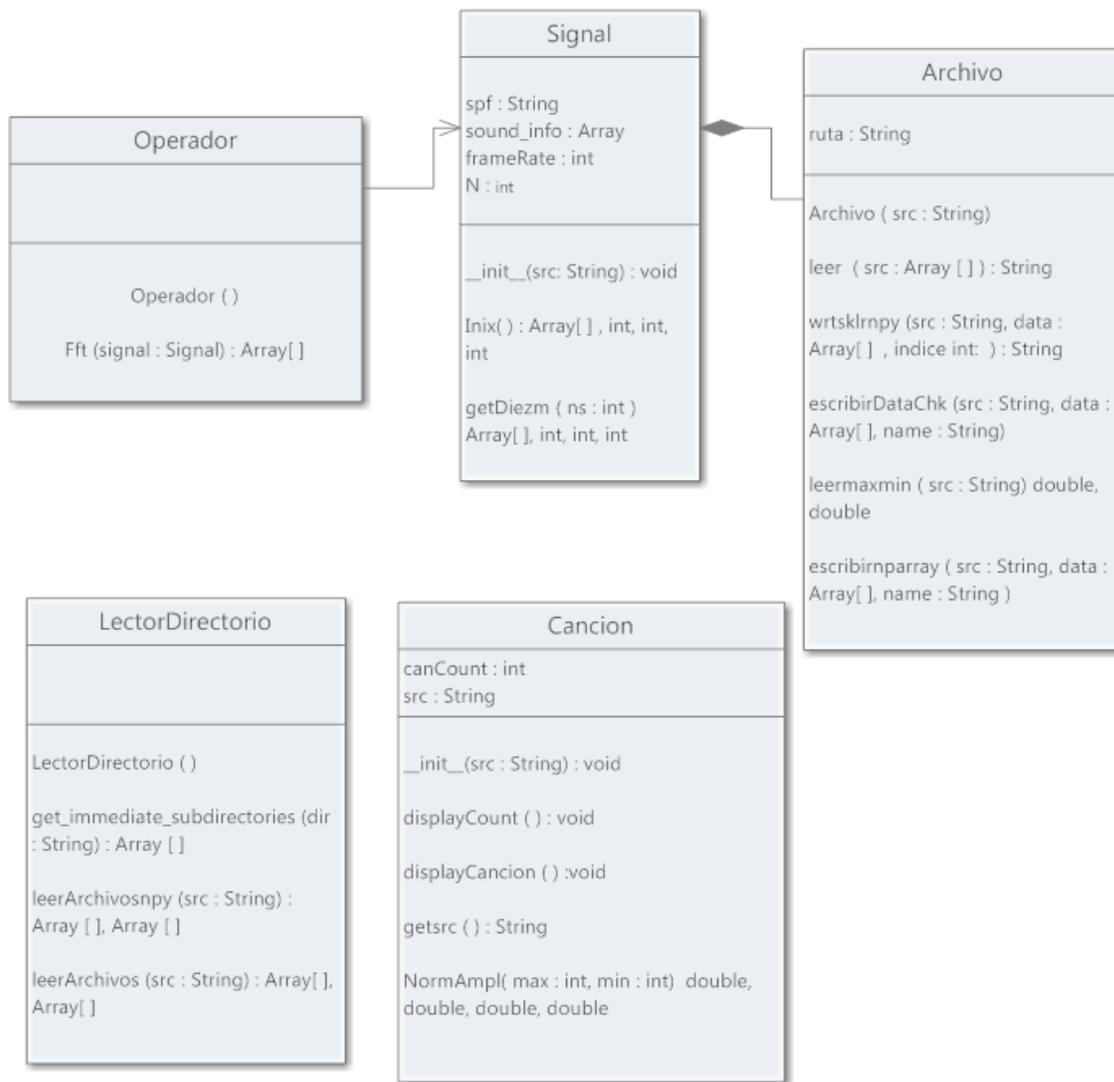
En el diagrama siguiente 7.8.1 se ilustran los casos de uso del Prototipo II versión 7.



**Imagen 7.8.1:** Diagrama de casos de uso del Prototipo II versión 7.

### 7.8.2 Diagrama de Clases

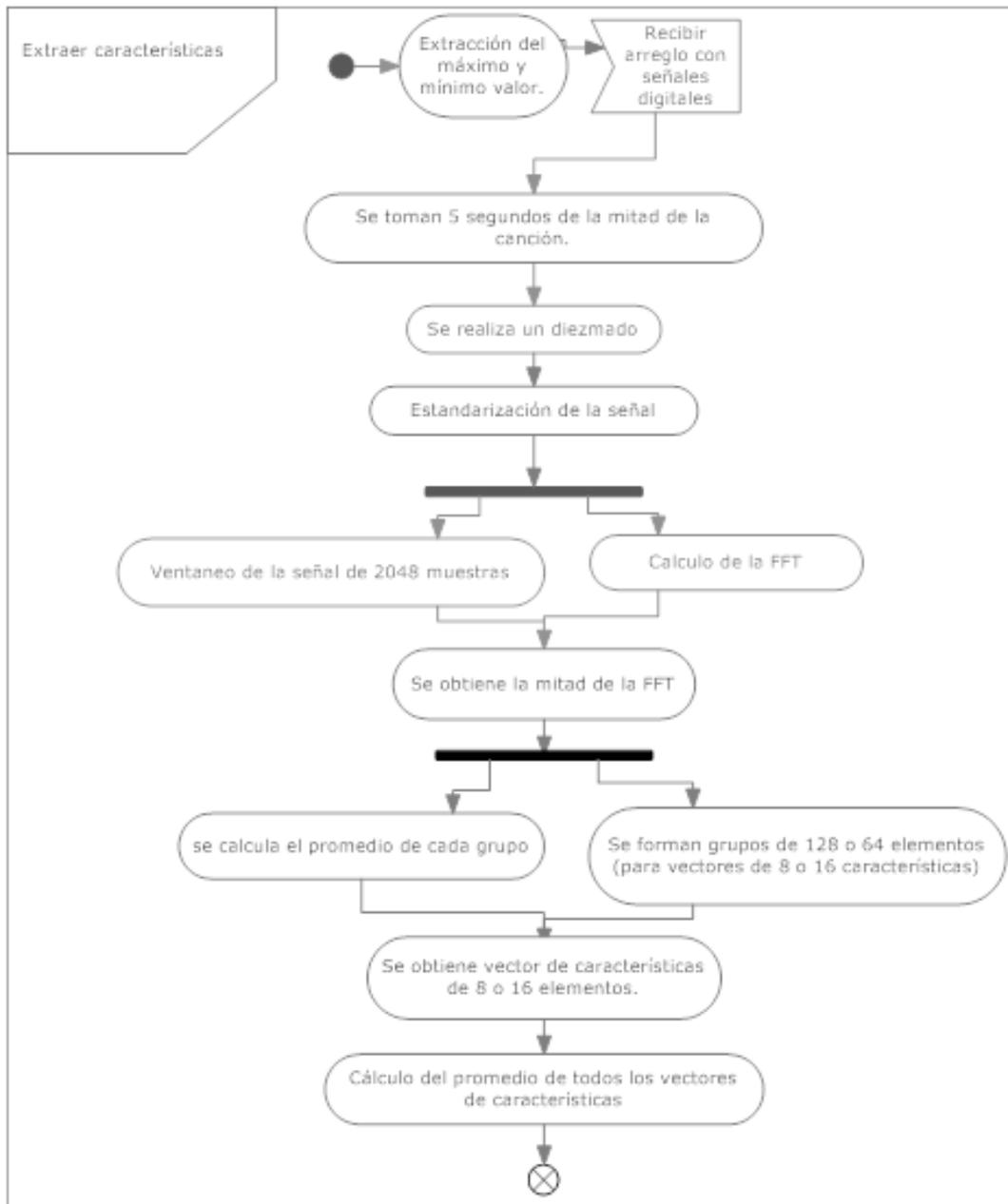
El siguiente diagrama 7.8.2 corresponde al diagrama de clases del Prototipo II versión 7.



**Imagen 7.8.2:** Diagrama de clases del Prototipo II versión 7.

### 7.8.3 Diagrama de Actividades

El siguiente diagrama 7.8.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 7.



**Imagen 7.8.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 7.

## 7.9 8° Versión

De acuerdo a los resultados obtenidos en la prueba en la que se muestra una gráfica con el promedio de la FFT de los siete géneros, observamos que en los primeros 50 elementos es en donde existe una mayor diferencia. Por este motivo se decide crear esta versión del Prototipo II.

Se toman los 5 segundos de la mitad de la canción y se toman ventanas de 2048 muestras y se calcula la FFT, después solo se toma la mitad es decir 1024 muestras por ventana posteriormente se promedian los elementos de cada una de las ventanas para así solo obtener una, de la que solo se toman los 50 primeros elementos que conforman nuestro vector de características por canción.

Para el entrenamiento se crea una sola muestra que es producto del promedio de todas las muestras participantes en el entrenamiento.

### 7.9.1 Requerimientos no funcionales

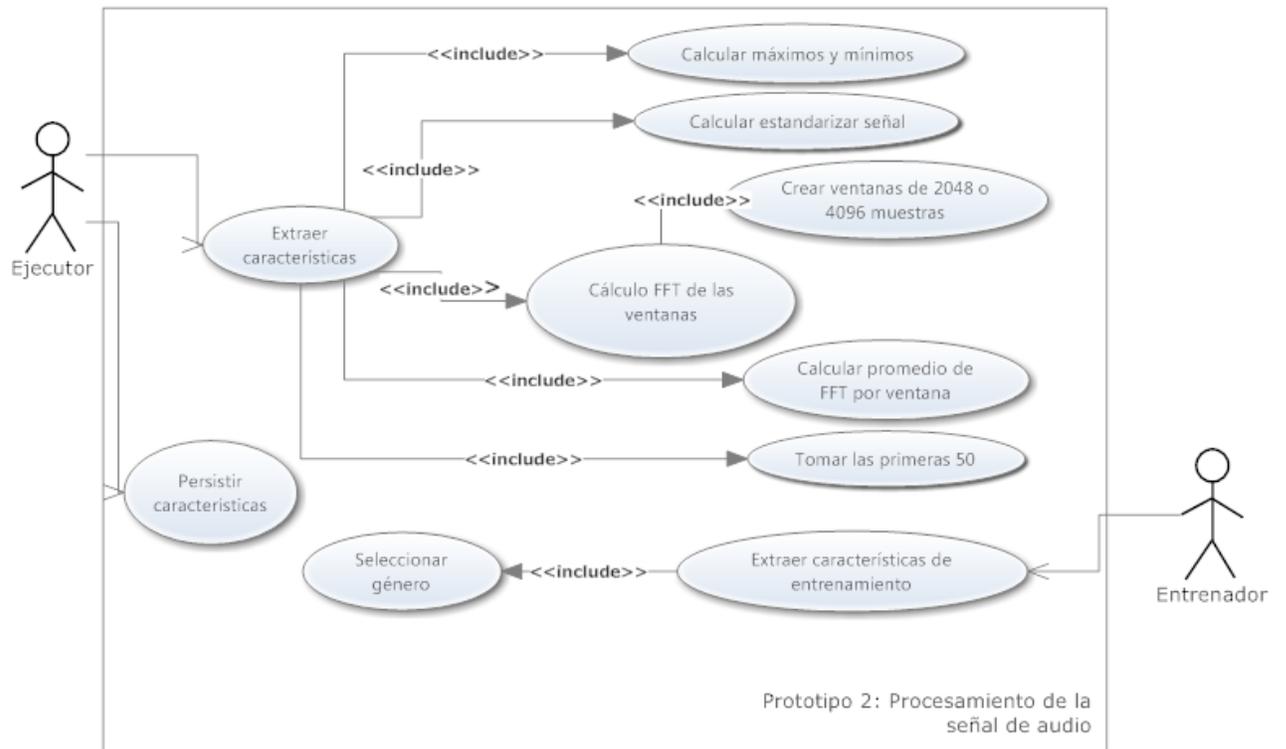
En la tabla 7.9.1, se enlistan los requerimientos no funcionales para el Prototipo

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de las señales de entrenamiento, ambos valores se almacenan en un archivo de texto plano llamado MaxMin.txt donde cada línea corresponde a un valor.
ReqNoF3	Se utilizan las muestras que corresponden a los 5 segundos de la mitad de cada canción para ser analizados. Considerando la velocidad de muestreo de los archivos WAV de 16,000 Hz, tenemos un total de 80,000 muestras.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Fourier de la señal por ventanas de 2048 muestras, después solo se toman las 1024 primeras muestras por la simetría de la otra mitad. Se hace un promedio del número de muestra $n$ de cada una de las ventanas, solo se toman las primeras 50 muestras para formar un vector de características que contiene el promedio de cada punto de la FFT de todas las ventanas.
ReqNoF6	El vector de características de cada canción es almacenado en un archivo de formato numpy en un archivo npy.
ReqNoF7	Para la etapa de pruebas se crea un solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta de nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF8	Para la etapa de entrenamiento se crea una carpeta con archivos en formato numpy. Al finalizar el proceso de entrenamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.
ReqNoF9	Para la etapa de entrenamiento se crea una sola muestra producto del promedio de todos los vectores de características de de las muestras involucradas en el entrenamiento.

**Tabla 7.9.1:** Tabla de requerimientos no funcionales del Prototipo II versión 8.

## 7.9.2 Diagrama de Casos de Uso

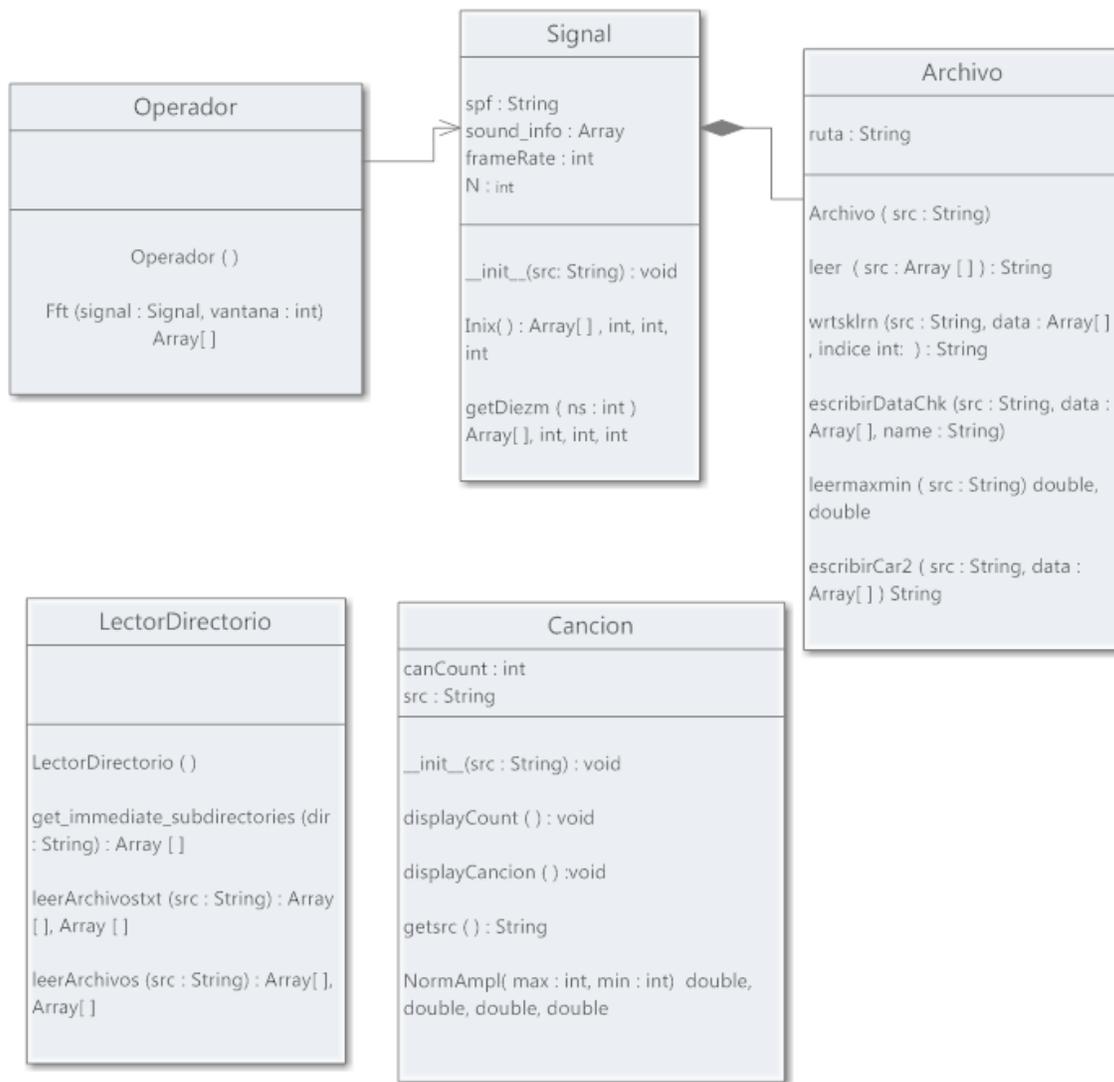
En el diagrama siguiente 7.9.1 se ilustran los casos de uso del Prototipo II versión 8.



**Imagen 7.9.1:** Diagrama de casos de uso del Prototipo II versión 8.

### 7.9.3 Diagrama de Clases

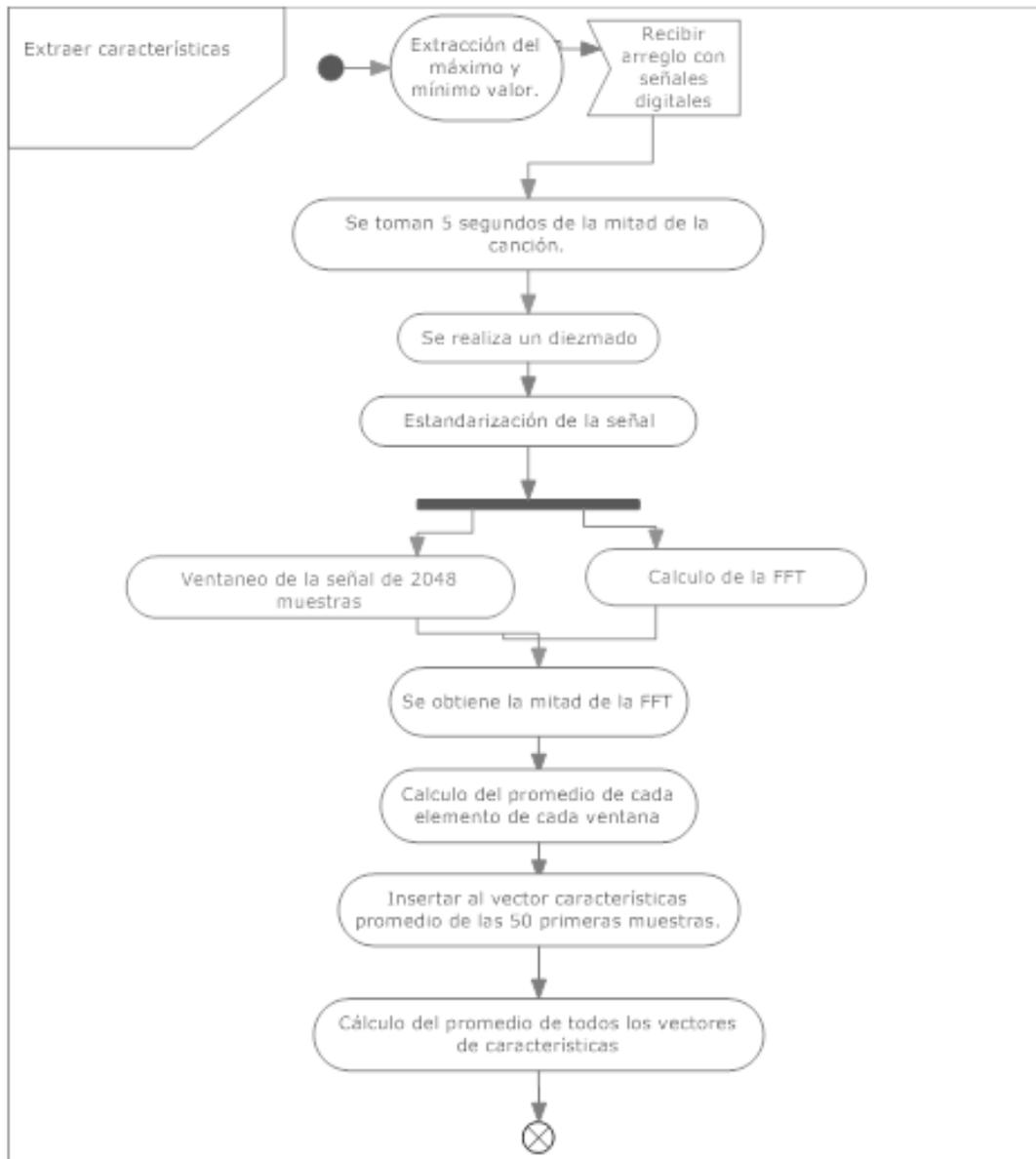
El siguiente diagrama 7.9.2 corresponde al diagrama de clases del Prototipo II versión 8.



**Imagen 7.9.2:** Diagrama de clases del Prototipo II versión 8.

#### 7.9.4 Diagrama de Actividades

El siguiente diagrama 7.9.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 8.



**Imagen 7.9.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 8.

## 7.10 9° Versión

Esta versión del Prototipo II tiene como finalidad experimentar con vectores de características que implementen la fase y el módulo resultante de la Transformada Rápida de Fourier.

Existen tres variantes de este prototipo. Cada una de estas variantes se describe a continuación.

- ⇒ La primer variante de este prototipo toma un fragmento de 10 segundos por la mitad de la canción y se realiza un ventaneo de 1024 muestras y se aplica la transformada rápida de Fourier por cada ventana y se conserva el módulo resultante. Se promedian los módulos resultantes de la FFT de todas las ventanas, obteniendo así un vector de características de 1024 elementos por cada canción. Para la etapa de entrenamiento se promedian todos los vectores de características de todas las muestras de entrenamiento, obteniendo así un solo vector de características por género entrenado.
- ⇒ Este prototipo toma los 5 segundos de la mitad de la canción y crea ventanas de 1024 elementos para después calcular la FFT y se obtiene la fase de cada elemento de la FFT. Se promedian todas las fases de todas las ventanas y se obtiene un vector de características de 1024 elementos. Para el entrenamiento se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.
- ⇒ Es prototipo calcula la fase y el módulo de la FFT de ventanas de 1024 características de un fragmento de la mitad de la canción de 5 segundos. Se obtiene un vector de 2048 elementos de los cuales 1024 elementos corresponden a la fase y 1024 al módulo. Para el entrenamiento se genera una sola referencia producto del promedio del vector de características de todas las muestras involucradas, así se tiene una sola referencia por género entrenado.

### 7.10.1 Requerimientos no funcionales

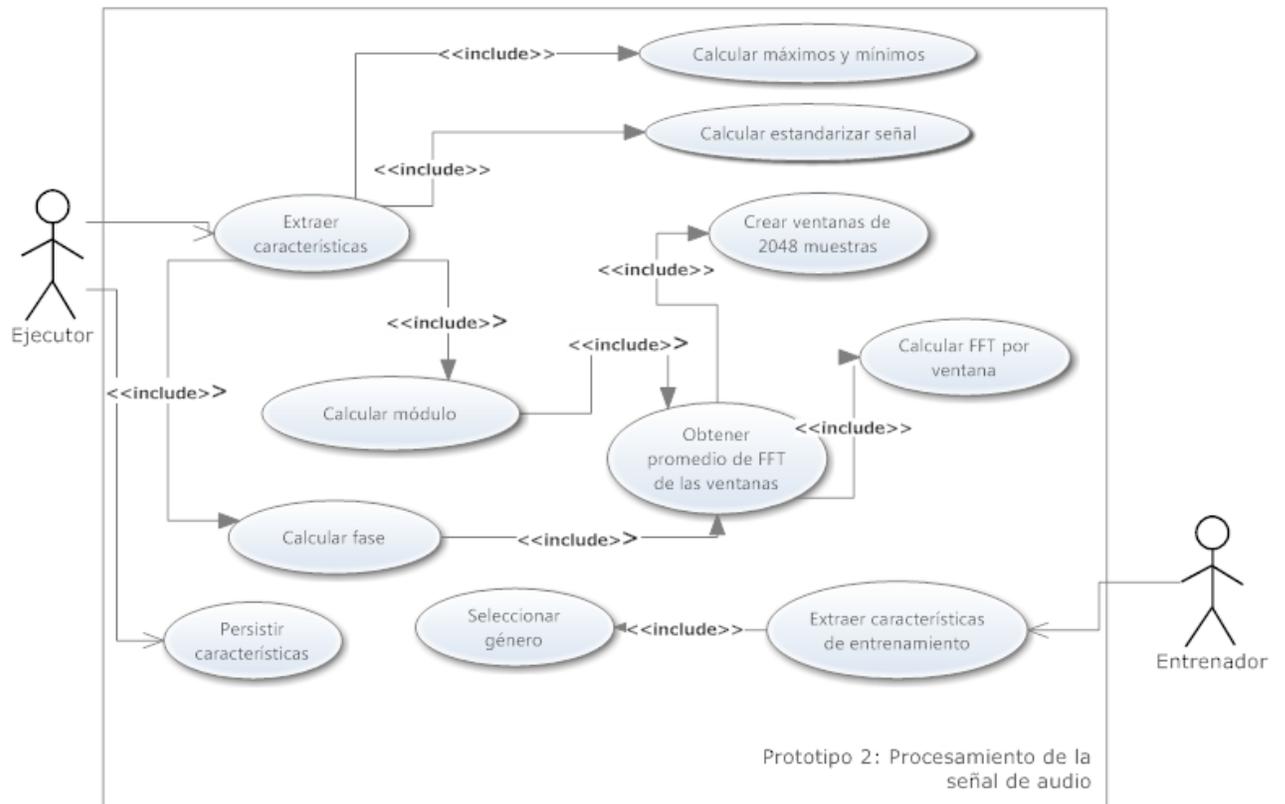
En la tabla 7.10.1, se enlistan los requerimientos no funcionales para el Prototipo

ID Requerimeinto	Requerimiento
ReqNoF1	Se recibirán las señales de audio en un arreglo bidimensional, donde cada fila corresponde a la señal de un archivo y la señal de cada archivo es un arreglo.
ReqNoF2	Se obtiene el valor máximo y el valor mínimo de las señales de entrenamiento, ambos valores se almacenan en un archivo de texto plano llamado MaxMin.txt donde cada línea corresponde a un valor.
ReqNoF3	Para la primer variante del prototipo se toman los 10 segundos de la mitad de la canción para ser analizados.
ReqNoF4	Para la segunda y tercer variante del prototipo se toman los 5 segundos de la mitad de la canción para ser analizados.
ReqNoF4	Se hace un diezmado del segmento de la señal a analizar si es que la velocidad de muestreo no es de 16,000 Hz y se estandarizan sus valores de acuerdo a los máximos y mínimos encontrados con anterioridad.
ReqNoF5	Se calcula la transformada rápida de Furier de la señal por ventanas de 2048 muestras, después solo se toman las 1024 primeras muestras por la simetría de la otra mitad. Se hace un promedio del número de muestra $n$ de cada una de las ventanas.
ReqNoF6	Para la primer variante se calcula el módulo de la ventana promedio y así se obtiene el vector de características.
ReqNoF6	Para la segunda variante se calcula la fase de la ventana promedio y así se obtiene el vector de características.
ReqNoF6	Para la tercer variante se calcula el modulo y la fase de la ventana promedio y así se obtiene el vector de características de 2048 elementos ya que son 1024 correspondiente al módulo y 1024 a la fase.
ReqNoF7	Para la etapa de pruebas se crea un solo archivo de texto llamado DataCheck.txt, donde cada línea contiene las características de la canción y su ruta absoluta. Este archivo se guarda dentro de una carpeta de nombre Test, que a su vez se encuentra contenida en la carpeta TestFFt.pyData.
ReqNoF8	Para la etapa de entrenamiento se crea una carpeta con archivos en formato numpy. Al finalizar el proceso de entranamiento se obtiene un archivo de texto con el nombre Data-trian.txt en la carpeta FFTPrincipal.pyData.
ReqNoF9	Para la etapa de entrenamiento se crea una sola muestra producto del promedio de todos los vectores de características de de las muestras involucradas en el entrenamiento.

**Tabla 7.10.1:** Tabla de requerimientos no funcionales del Prototipo II versión 9.

### 7.10.2 Diagrama de Casos de Uso

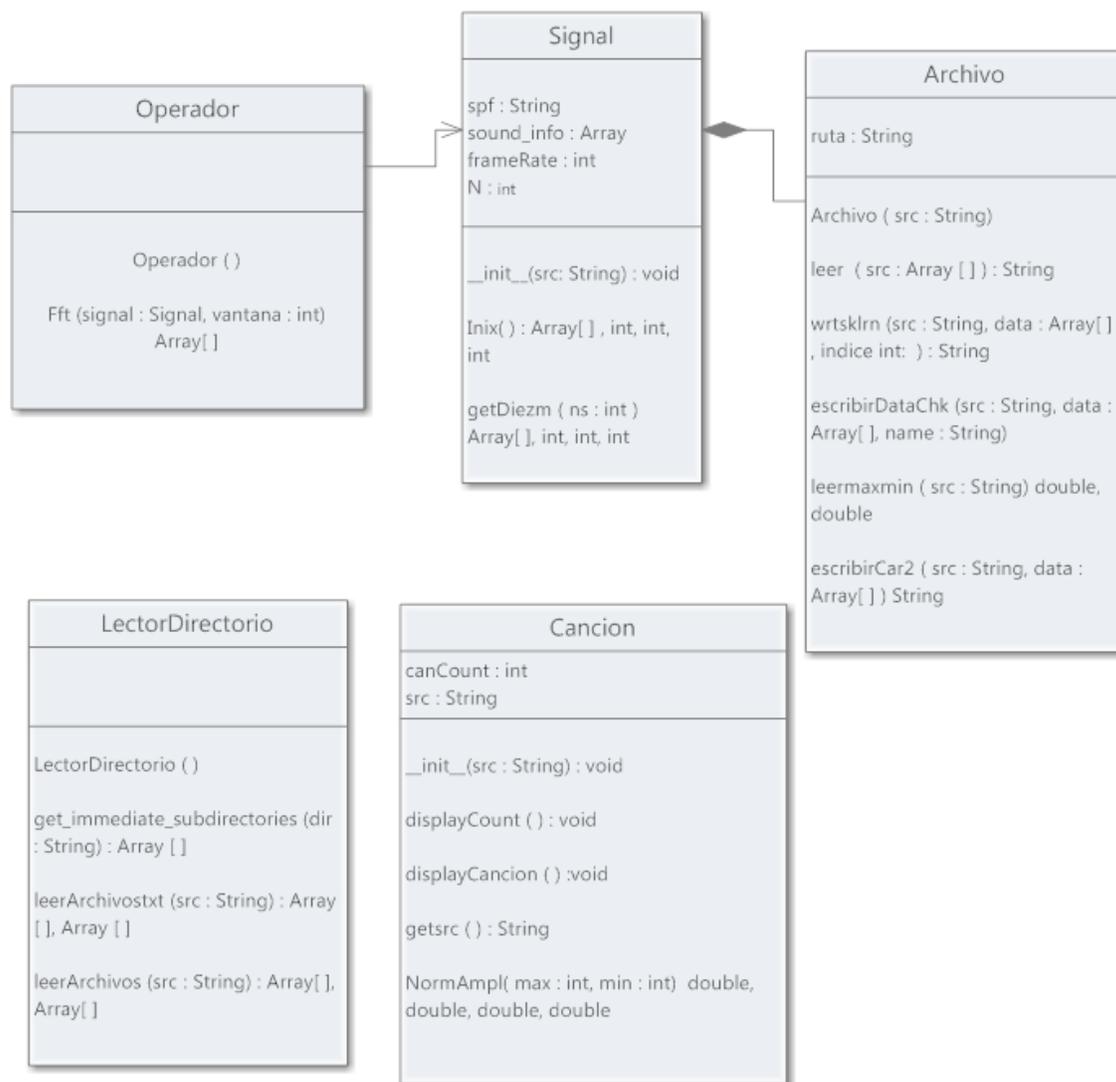
En el diagrama siguiente 7.10.1 se ilustran los casos de uso del Prototipo II versión 9.



**Imagen 7.10.1:** Diagrama de casos de uso del Prototipo II versión 9.

### 7.10.3 Diagrama de Clases

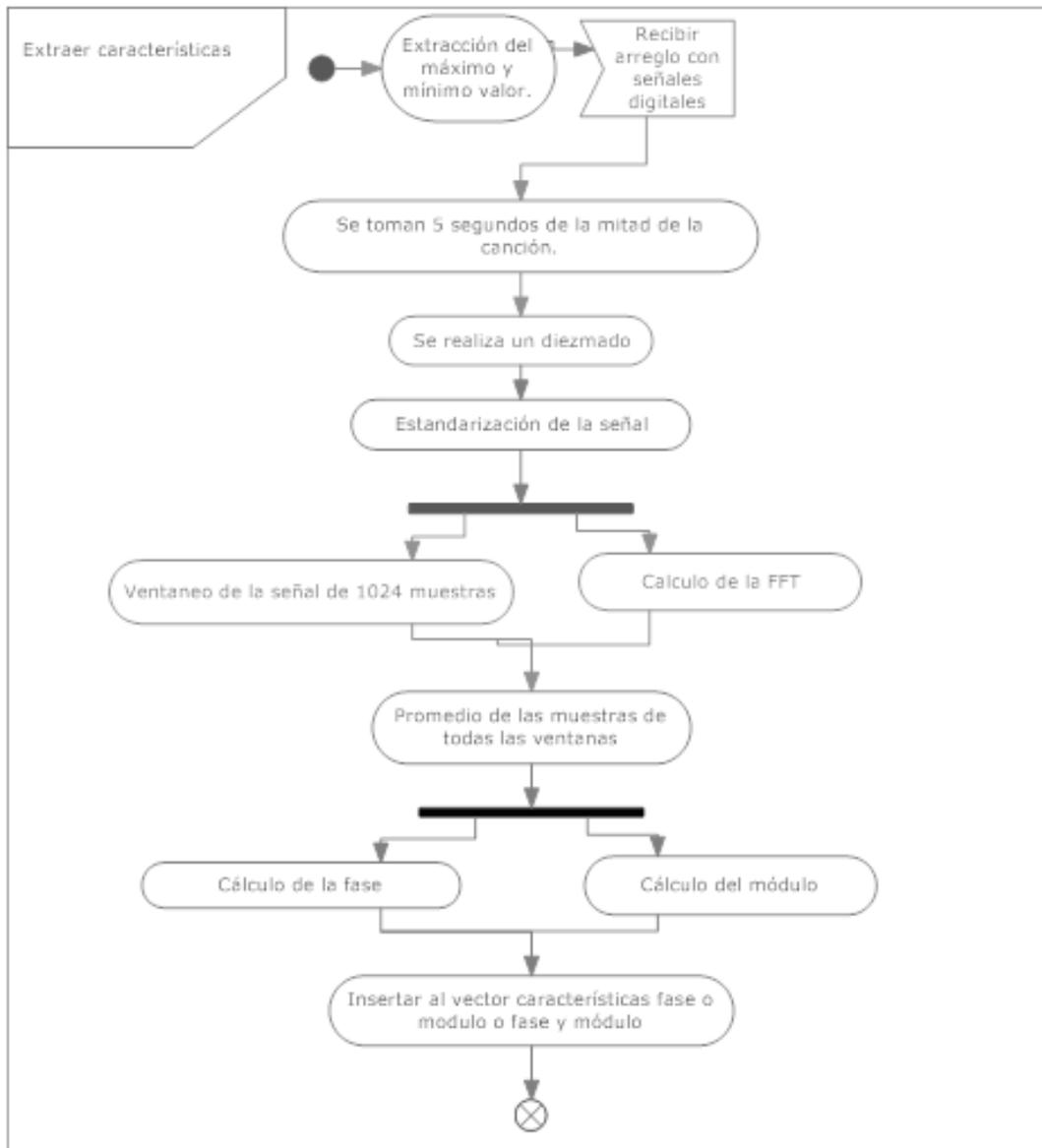
El siguiente diagrama 7.10.2 corresponde al diagrama de clases del Prototipo II versión 9.



**Imagen 7.10.2:** Diagrama de clases del Prototipo II versión 9.

#### 7.10.4 Diagrama de Actividades

El siguiente diagrama 7.10.3, corresponde al diagrama de actividades de la extracción de características del Prototipo II versión 9.



**Imagen 7.10.3:** Diagrama de actividades de la extracción de características del Prototipo II versión 9.

# 8

## Prototipo III: Clasificador de archivos musicales

### 8.1 1º Versión: Vecinos más cercanos.

Esta versión del prototipo utiliza un clasificador supervisado del vecino más cercano. El principio básico de este algoritmo de clasificador es el de encontrar un número predefinido de muestras de entrenamiento más cercanas en distancia a la nueva muestra, y predecir la clase a la cual pertenece. El número de muestras de entrenamiento cercanas a la muestra de prueba, puede ser definido por el usuario como una constante durante el proceso, o puede variar dependiendo de la densidad de muestras cercanas.

Esta versión 7.5.1 del prototipo hace uso de la clase `KNeighborsClassifier` que nos proporciona la librería `neighbors` de `scikit-learn` [13]. El prototipo crea una instancia de la clase

KNeighborsClassifier con su constructor, enviando como parámetros `algorithm`, `leaf_size`, `n_neighbors`, `p`, `warn_on_equidistant` y `weights`.

- ⇒ El parámetro `algorithm`, recibe el algoritmo que se empleará para la clasificación, este parámetro puede recibir los valores `auto`, `ball_tree`, `kd_tree` o `brute`, cuando nosotros instanciamos nuestro clasificador lo hacemos con `auto`, así se selecciona automáticamente el mejor algoritmo para la clasificación.
- ⇒ El parámetro `leaf_size`. Entre más grande sea este valor se construyen más rápido los árboles de decisión, porque el número de nodos que necesitan crearse se reducen. El valor por defecto es 30.
- ⇒ El parámetro `n_neighbors` indica el número de vecinos que se toman en cuenta para tomar la decisión.
- ⇒ El parámetro `weights`. La clasificación básica de los vecinos más cercanos utiliza ponderaciones uniformes, es decir, el valor asignado a un punto de consulta se calcula a partir de una mayoría simple de los vecinos más cercanos. Cuando se asigna "uniform" asigna ponderaciones uniformes para cada vecino. Cuando se asigna "distance" se dan pesos proporcionales a la inversa de la distancia desde el punto de consulta. [13].

En la siguiente figura 8.1.1 se puede observar el diagrama de la clase que integra a la 1ª versión del Prototipo III.



**Imagen 8.1.1:** Diagrama de clases del Prototipo III KNN.

## 8.2 2º Versión: Máquina de soporte vectorial.

El clasificador usado por esta versión del prototipo III es una máquina de soporte vectorial del tipo SVC por sus siglas en inglés Support Vector Clustering con Funcion de Base Radial RBF por sus siglas en inglés (Radial Basis Function).

Se hace uso de la librería scikit-learn [13], esta librería nos proporciona la clase svm con la que se crea una instancia con el método SVC(), que recibe el parámetro gamma, mismo que variamos en las pruebas para ajustar el clasificador y obtener mejores resultados.

- ⇒ El parámetro C es común a todas las SVM, opera fuera de los errores de clasificación contra la superficie de decisión. Un valor pequeño de C hace que la decisión sobre la superficie sea menos estricta, mientras que un alto valor en C tiene como objetivo una clasificación más estricta.
- ⇒ El parámetro gamma define la influencia que tiene una muestra de entrenamiento. El

valor de gamma que tanto afecta la cercanía de las muestras de entrenamiento.

En la siguiente figura 8.2.1 se puede observar el diagrama de la clase que integra a la 2º versión del Prototipo III.



**Imagen 8.2.1:** Diagrama de clases del Prototipo III SVM.

# 9

## Prototipo IV: Prototipo Integrador

La función de este prototipo es el de presentar una interfaz gráfica amigable y de uso sencillo que integre la funcionalidad de los prototipos I, II y III con las versiones que tienen mejores resultados.

Este prototipo debe permitir al usuario entrenar y realizar la predicción de género de alguna pista musical seleccionada por el usuario.

### 9.1 Requerimientos funcionales genéricos

En la tabla 9.1.1 se enlistan los requerimientos funcionales genéricos para cualquiera de las versiones de este Prototipo IV.

ID Requerimeinto	Requerimiento
ReqNoF1	Se presenta una interfaz gráfica.
ReqNoF2	Permite realizar el entrenamiento y seleccionar los archivos musicales que se utilizaran.
ReqNoF3	Permite realizar la predicción de género de archivos musicales.

**Tabla 9.1.1:** Tabla de requerimientos funcionales genéricos del Prototipo IV.

## 9.2 1° Versión.

### 9.2.1 Requerimientos no funcionales

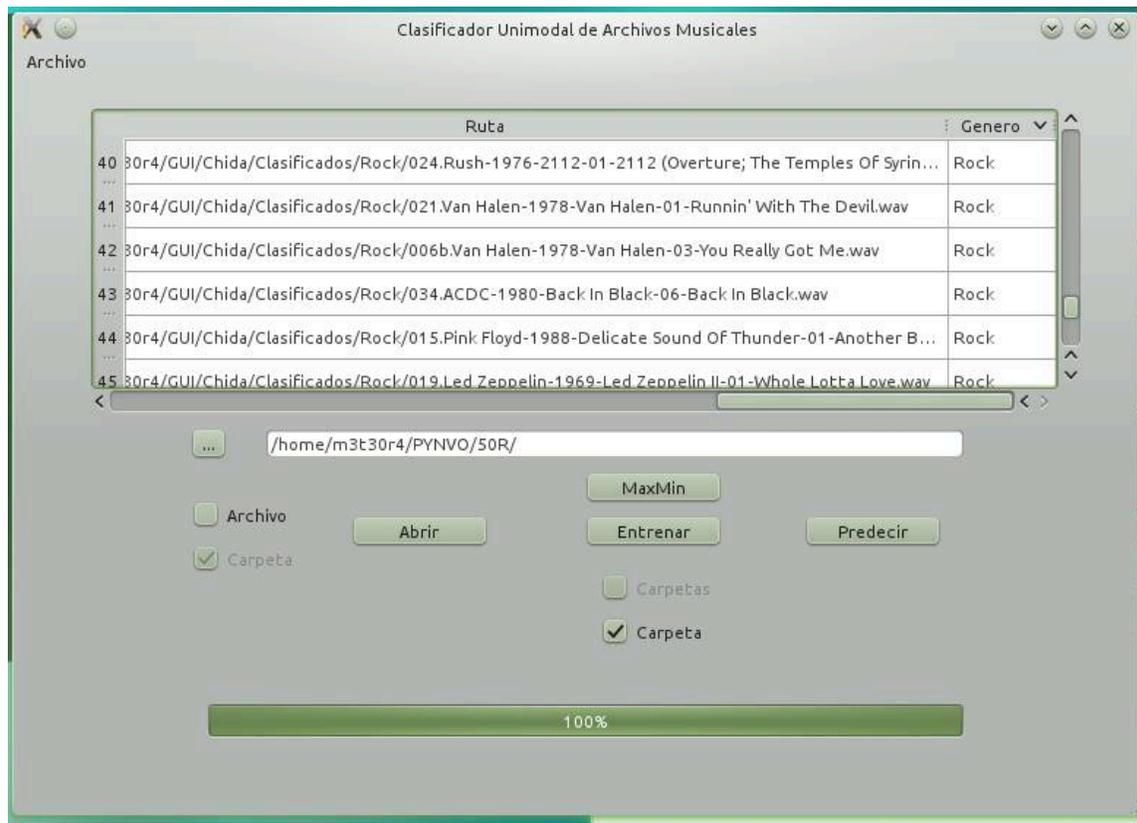
En la tabla 9.2.1 se enlistan los requerimientos no funcionales para la primer versión del Prototipo IV.

ID Requerimeinto	Requerimiento
ReqNoF1	Para realizar el entrenamiento se debe presionar el boton con la leyenda Entrenar.
ReqNoF2	Para predecir existe un botón con la leyenda Predecir.
ReqNoF3	Se cuenta con un área de texto para ingresar la ruta del archivo o del directorio con los archivos que el usuario seleccione, también se cuenta con una botón a un lado de esta área para abrir un explorador de archivos y así buscar el directorio o archivo, al seleccionar el deseado, se mostrará en la barra de texto.
ReqNoF4	Para saber si la selección realizada por el usuario es un archivo o un directorio se cuenta con dos checkbox identificados por la leyenda Carpeta y Archivo respectivamente. Estos deberán ser seleccionados por el usuario después de su selección de archivo o directorio.
ReqNoF5	Se muestra una barra de progreso que indica el avance del procesamiento de los archivos.
ReqNoF6	Para realizar la predicción solo se permite un archivo.
ReqNoF7	Los directorios seleccionados por el usuario deberán contener archivos en formato WAV con la extensión wav, de lo contrario los demás archivos que se encuentren en ese directorio serán ignorados.
ReqNoF8	Los archivos WAV deberan ser de una velocidad de 16,000 Hz, con una calidad mono a 16 bits.

**Tabla 9.2.1:** Tabla de requerimientos no funcionales de la versión 1 del Prototipo 2.

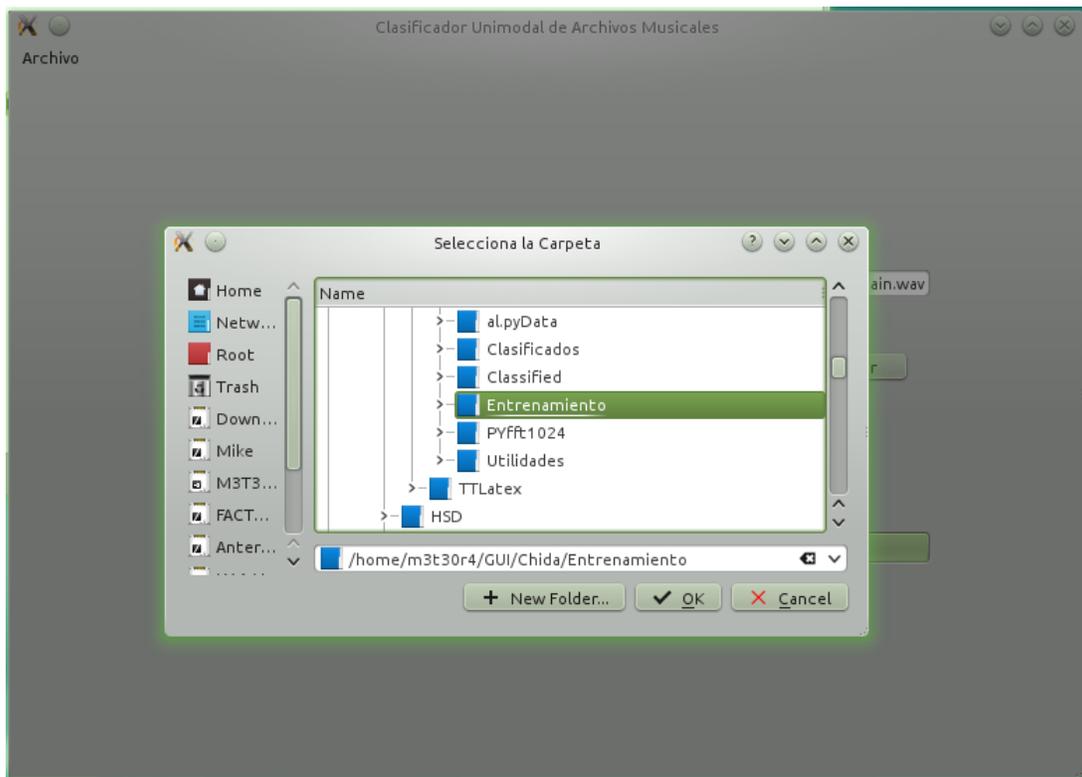
### 9.2.2 Diseño de la interfaz

En la imagen 9.2.1 se observa la pantalla principal de la interfaz gráfica del Prototipo IV.



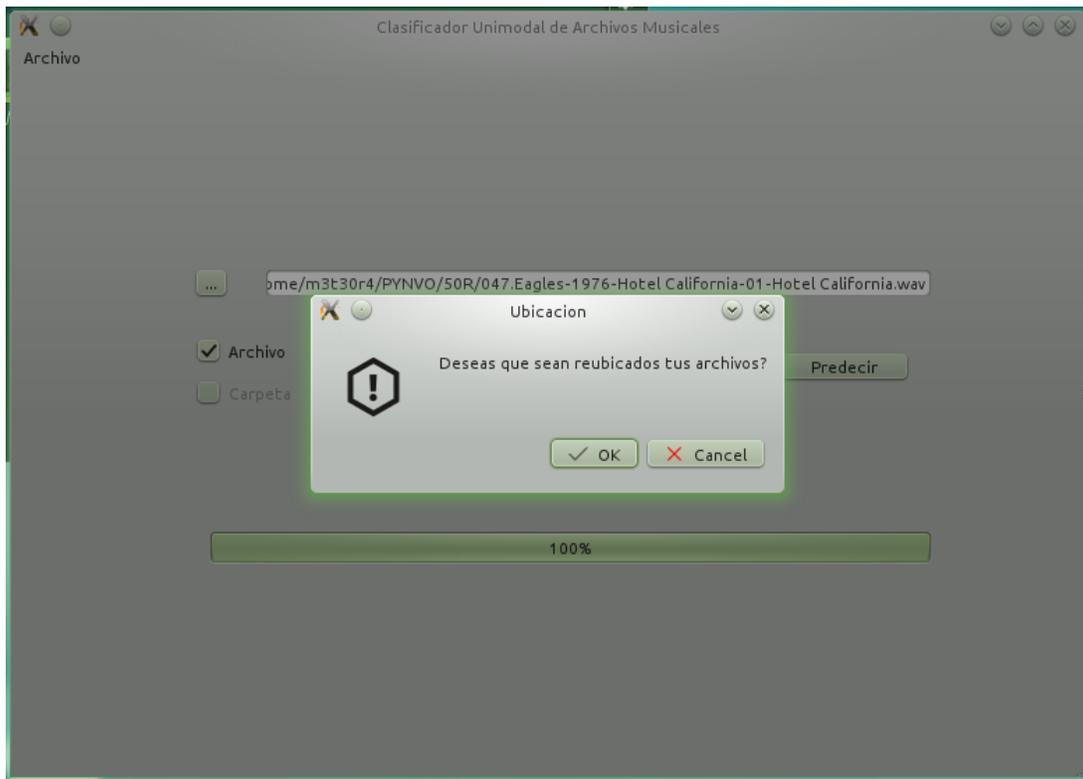
**Imagen 9.2.1:** Diseño de la pantalla principal del Prototipo IV versión 1.

En la imagen 9.2.2 se observa la ventana emergente para seleccionar el directorio del género que se entrena.



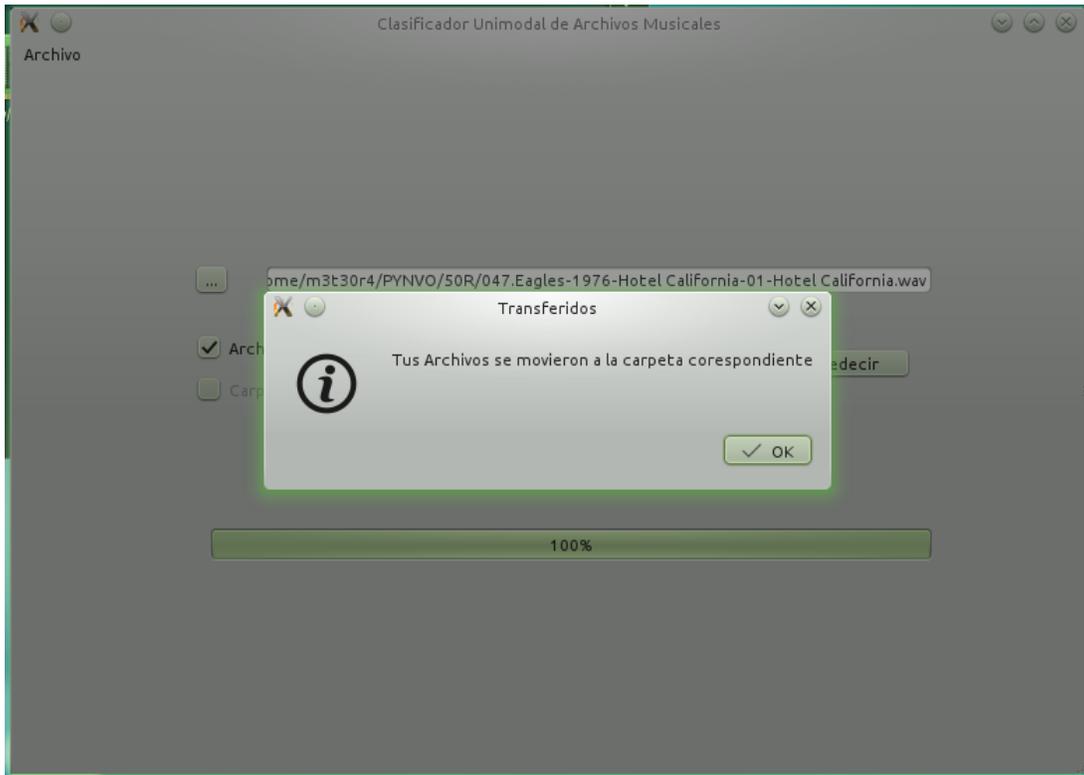
**Imagen 9.2.2:** Mensaje para seleccionar el género a entrenar.

En la imagen 9.2.3 se observa el mensaje que indica que se ha clasificado.



**Imagen 9.2.3:** Mensaje que muestra que se ha clasificado.

En la imagen 9.2.4 se observa un mensaje para indicar que los archivos han sido transferidos.



**Imagen 9.2.4:** Mensaje que muestra que se han transferido los archivos.

# 10

## Pruebas de clasificación

Las pruebas mostradas a continuación fueron realizadas con archivos en formato WAV a una frecuencia de muestreo de 16,000 Hz, con una codificación de 16 bits y con canal mono.

### 10.1 1° Prueba

#### 10.1.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

1. Prototipo I: 1.
2. Prototipo II: 1.

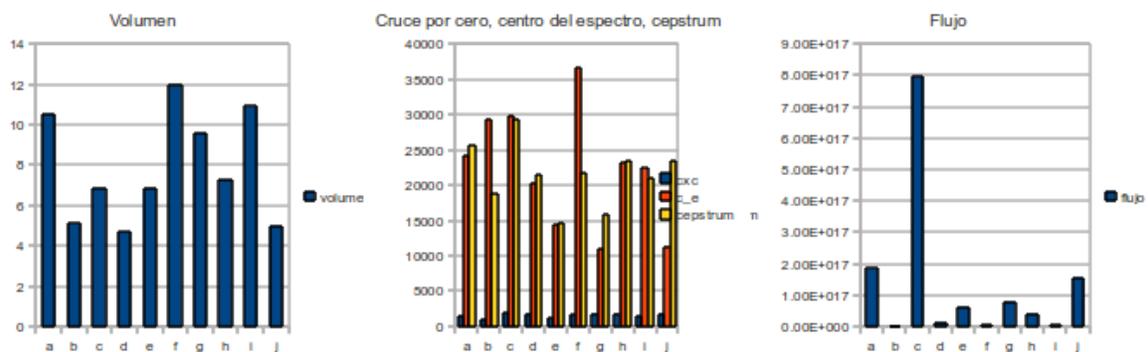
Esta prueba se realizó con la finalidad de poder observar de manera visual las variaciones entre las características obtenidas por la primer versión del prototipo II.

### 10.1.2 Resultados

En las imágenes siguientes se muestran las gráficas de los valores obtenidos de las características volumen, cruce por cero, centro del espectro, cepstrum y flujo.

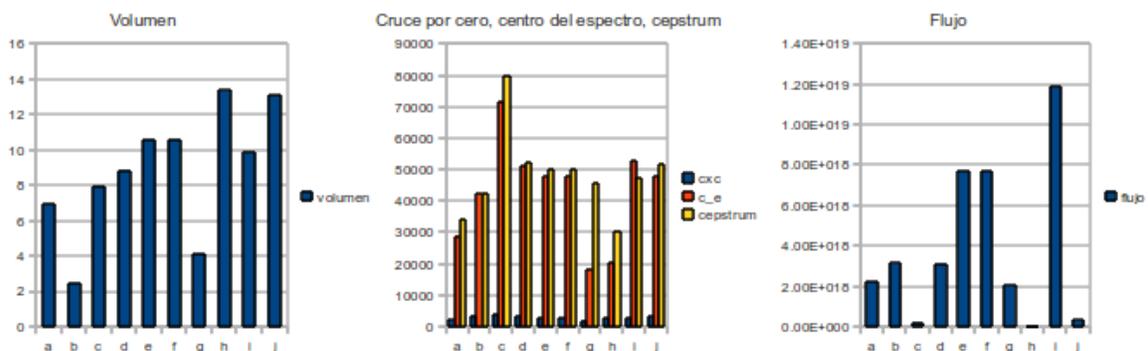
Para realizar estas gráficas se tomó una muestra de 10 canciones por cada género.

En las siguientes gráficas 10.1.1 se observan los resultados para el género clásica.



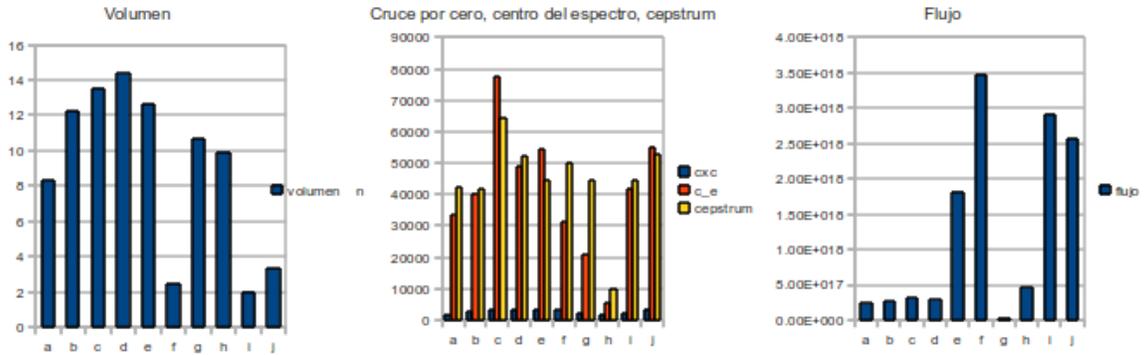
**Imagen 10.1.1:** Características en el tiempo del género clásica.

En las siguientes gráficas 10.1.2 se observan los resultados para el género electrónica.



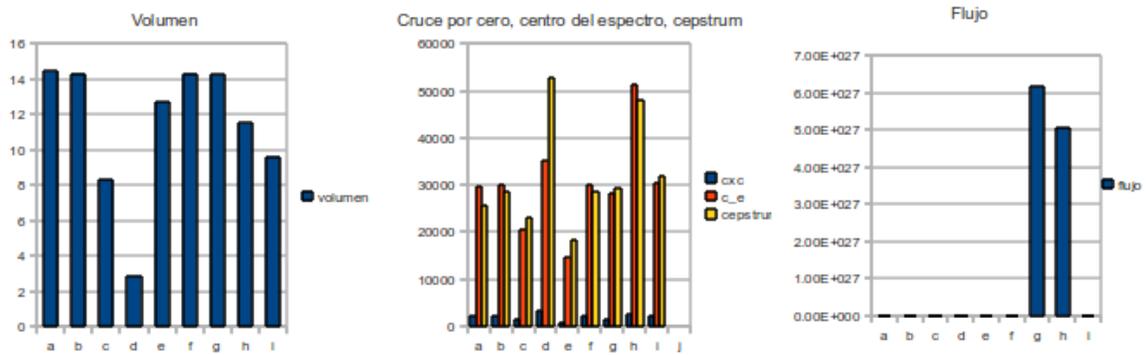
**Imagen 10.1.2:** Características en el tiempo del género electrónica.

En las siguientes gráficas 10.1.3 se observan los resultados para el género hip-hop.



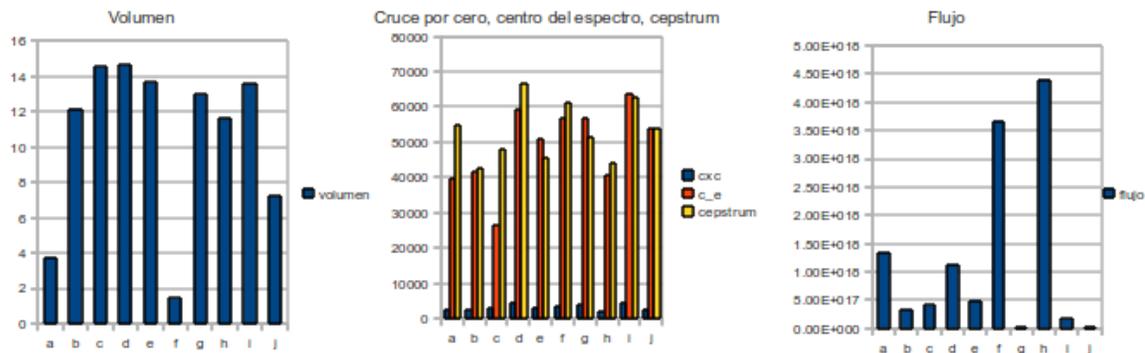
**Imagen 10.1.3:** Características en el tiempo del género hip-hop.

En las siguientes gráficas 10.1.4 se observan los resultados para el género jazz.



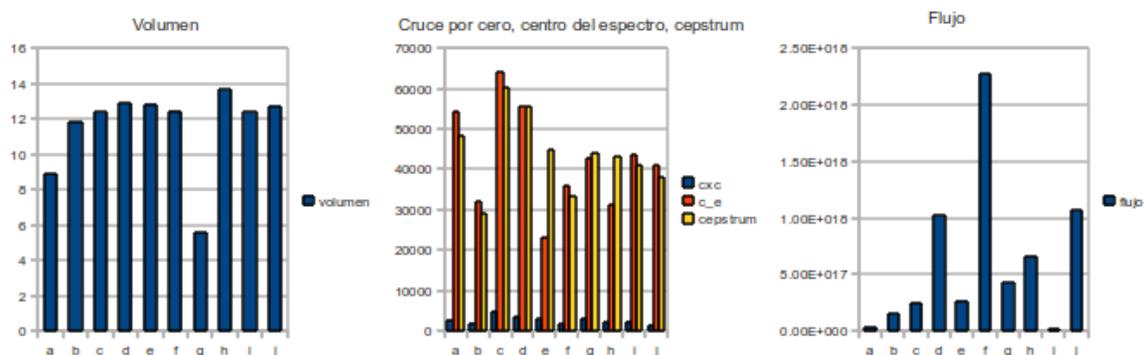
**Imagen 10.1.4:** Características en el tiempo del género jazz.

En las siguientes gráficas 10.1.5 se observan los resultados para el género pop.



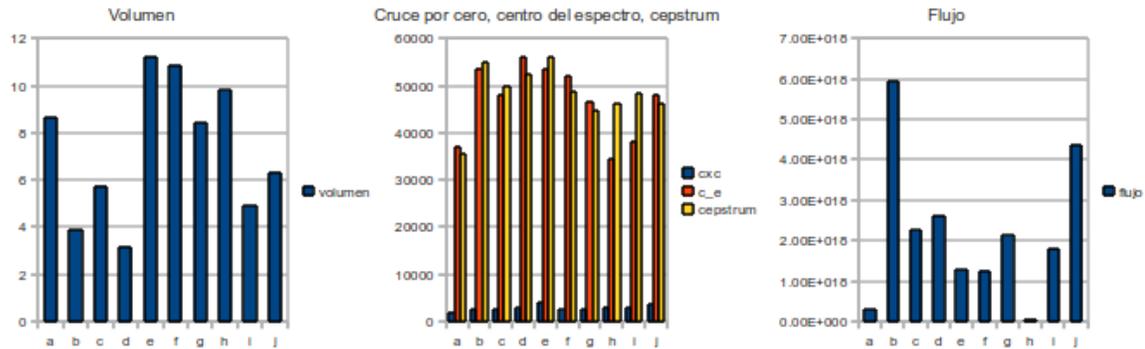
**Imagen 10.1.5:** Características en el tiempo del género pop.

En las siguientes gráficas 10.1.6 se observan los resultados para el género rock.



**Imagen 10.1.6:** Características en el tiempo del género rock.

En las siguientes gráficas 10.1.7 se observan los resultados para el género salsa.



**Imagen 10.1.7:** Características en el tiempo del género salsa.

## 10.2 2° Prueba

### 10.2.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

1. Prototipo I: 1.
2. Prototipo II: 1.
3. Prototipo III: 1 y 2.

Se entrenó el prototipo con los siete géneros: clásica, electrónica, hip-hop, jazz, pop, rock, salsa. Las pruebas fueron hechas en 7 etapas, donde en cada etapa se prueban 10 canciones del mismo género. Las características que se toman en cuenta son en el dominio del tiempo.

Las características obtenidas son el cruce por cero, tono, flujo, roll off, volumen y centro del espectro.

### 10.2.2 Resultados

Los resultados de las pruebas realizadas fueron muy pobres. El clasificador que emplea una máquina de soporte vectorial arroja el género de hip-hop, mientras que el clasificador KNN

devuelve como resultado el género clásica cuando se hace la prueba con cada uno de los 7 géneros de entrenamiento.

En la siguiente matriz 10.2.1 se muestran las predicciones obtenidas por cada género probado con cada uno de los clasificadores.

			Rock	Pop	Jazz	Clásica	Salsa	Hiphop	Electrónica
EXPECTED	Rock	SVM						100	
		KNN				100			
	Pop	SVM						100	
		KNN				100			
	Jazz	SVM						100	
		KNN				100			
	Clásica	SVM							100
		KNN				100			
	Salsa	SVM							100
		KNN				100			
	Hip-hop	SVM							100
		KNN				100			
	Electrónica	SVM							100
		KNN				100			

**Imagen 10.2.1:** Matriz de resultados prueba 2.

### 10.2.3 Conclusiones

Por los resultados se deduce que las características obtenidas no son óptimas para clasificar, pues quedan distribuidos por todos el espacio de los clasificadores muestras de entrenamiento del género clásica para el clasificador KNN y hip-hop para el clasificador SVM.

## 10.3 3° Prueba

### 10.3.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

- ⇒ Prototipo I: 1.
- ⇒ Prototipo II: 2.
- ⇒ Prototipo III: 1 y 2.

De igual forma que en la prueba anterior se realizó la etapa de entrenamiento con los 7 géneros. Las pruebas fueron hechas en 7 etapas, donde en cada etapa se prueban 10 canciones del mismo género. La diferencia en los resultados es producto de las modificaciones al prototipo 2 que seleccionan 5 segundos de la mitad de la canción y estandarizan los valores de la señal de acuerdo a los valores máximos y mínimos de las señales de las canciones de los archivos de entrenamiento.

### 10.3.2 Resultados

Los resultados fueron muy parecidos a los de la prueba anterior, el Prototipo III versión 1 arroja siempre como predicción el género pop, mientras que la versión 2 muestra siempre como predicción el género de jazz.

En la siguiente matriz 10.3.1 se muestran las predicciones obtenidas por cada género probado con cada uno de los clasificadores.

			Rock	Pop	Jazz	Clásica	Salsa	Hiphop	Electrónica
EXPECTED	Rock	SVM		100					
		KNN			100				
	Pop	SVM		100					
		KNN			100				
	Jazz	SVM		100					
		KNN			100				
	Clásica	SVM		100					
		KNN			100				
	Salsa	SVM		100					
		KNN			100				
	Hip-hop	SVM		100					
		KNN			100				
	Electrónica	SVM		100					
		KNN			100				

Imagen 10.3.1: Matriz de resultados prueba 3.

## 10.4 4° Prueba

### 10.4.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

⇒ Prototipo I: 1.

⇒ Prototipo II: 3.

⇒ Prototipo III: 1 y 2.

De igual forma que en la prueba anterior se realizó la etapa de entrenamiento con los 7 géneros. Las pruebas fueron hechas en 7 etapas, donde en cada etapa se prueban 10 canciones del mismo género.

#### 10.4.2 Resultados

La versión 2 del Prototipo III arroja siempre la predicción de jazz, sin importar el género del conjunto de datos introducido. Mientras que la versión 1 del Prototipo III (KNN) alcanzó un mejor desempeño aunque no óptimo, ya que su predicción fue variable dependiendo de las canciones de cada conjunto de entrenamiento (cada conjunto de entrenamiento está compuesto por 10 canciones de un mismo género), obteniendo así su mejor resultado probando el género Jazz del cual logró clasificar correctamente el 50% (5 canciones) correctamente.

En la siguiente matriz 10.4.1 se muestran las predicciones obtenidas por cada género probado con cada uno de los clasificadores.

			Rock	Pop	Jazz	Clásica	Salsa	Hiphop	Electrónica
EXPECTED	Rock	SVM			100				
		KNN	60		30	10			
	Pop	SVM			100				
		KNN	40		10	40		10	
	Jazz	SVM			100				
		KNN	10		50	40			
	Clásica	SVM			100				
		KNN		20	20	10	30	10	10
	Salsa	SVM			100				
		KNN			10	60		20	10
	Hip-hop	SVM			100				
		KNN	70		10	10		10	
	Electrónica	SVM			100				
		KNN	50		10	10		30	

Imagen 10.4.1: Matriz de resultados prueba 4.

## 10.5 5° Prueba

### 10.5.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

⇒ Prototipo I: 1.

⇒ Prototipo II: 3.

⇒ Prototipo III: 1 y 2.

Se formaron grupos de dos géneros con los 7 géneros que se han tomado en cuenta (clásica, rock, pop, hiphop, salsa, electrónica y jazz). Para cada par de géneros se realizó un entrenamiento, y en base a este entrenamiento se realizan dos pruebas (una para cada género) con un conjunto de 10 canciones de cada uno de los géneros con los que se entrenó.

### 10.5.2 Resultados

En la imagen 10.5.1 se muestra una matriz que contiene los resultados de esta prueba. La primer columna corresponde al género que se espera obtener, y las intersecciones con las siguientes columnas muestran los resultados obtenidos al entrenarse con esos géneros y con cada versión del Prototipo III, representados en la tabla por SVM.

**Imagen 10.5.1:** Tabla que muestra los resultados de la prueba con SVM prueba 5.

En la siguiente gráfica 10.5.2 se muestra una matriz que contiene los resultados de esta prueba. La primer columna corresponde al género que se espera obtener, y las intersecciones con las siguientes columnas muestran los resultados obtenidos al entrenarse con esos géneros y con cada versión del Prototipo III, representados en la tabla por KNN.

	Rock	Pop	Jazz	Clasica	Salsa	Hiphop	Electronica
	KNN	KNN	KNN	KNN	KNN	KNN	KNN
Rock	40	30	10	0	10	40	30
Pop	40	80	0	0	0	80	70
Jazz	100	80	70	80	70	70	80
Clasica	70	70	80	80	80	30	40
Salsa	80	100	40	0	80	100	100
Hiphop	20	20	0	0	0	80	50
Electronica	50	80	10	80	10	80	80

**Imagen 10.5.2:** Tabla que muestra los resultados de la prueba con KNN prueba 5.

## 10.6 6° Prueba

### 10.6.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

- ⇒ Prototipo I: 1.
- ⇒ Prototipo II: 4.
- ⇒ Prototipo III: 1 y 2.

Para esta prueba se generaron pares con las combinaciones de géneros. Para cada combinación se realizó un entrenamiento con los dos géneros del par correspondiente y se realizaron 2 pruebas de un conjunto de 10 canciones cada una, donde cada conjunto pasee canciones de un solo género perteneciente a uno de los géneros de entrenamiento.

### 10.6.2 Resultados

En la siguiente matriz 10.6.1 se muestran las predicciones obtenidas por cada género probado con cada uno de los clasificadores.

	Rock		Pop		Jazz		Clasica		Salsa		Hiphop		Electronica	
	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN
Rock			100	80	100	10	80	0	100	80	70	40	0	90
Pop	0	40			100	20	100	20	0	70	80	80	50	40
Jazz	10	80	0	30			50	10	20	40	50	40	50	40
Clasica	10	70	0	30	0	80			100	0	0	30	0	80
Salsa	0	20	100	0	70	0	0	60			90	20	100	30
hiphop	0	50	0	40	0	80	60	20	0	90			60	80
Electronica	100	90	30	80	20	30	80	90	0	40	60	30		

Imagen 10.6.1: Matriz de resultados prueba 6.

## 10.7 7° Prueba

### 10.7.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

- ⇒ Prototipo I: 1.
- ⇒ Prototipo II: 5.
- ⇒ Prototipo III: 1 y 2.

En esta prueba se realizó un único entrenamiento que contemplaba los 7 géneros con los se ha trabajado. Se realizaron 7 pruebas cada una con un conjunto de 10 canciones de un solo género.

### 10.7.2 Resultados

Los resultados no fueron óptimos, en la mayoría de las pruebas el clasificador de la máquina de soporte vectorial devuelve como predicción el género pop, mientras que el clasificador del KNN devuelve como predicción el género electrónica.

En la siguiente matriz 10.7.1 se muestran las predicciones obtenidas por cada género probado con cada uno de los clasificadores.

			Rock	Pop	Jazz	Clásica	Salsa	Hiphop	Electrónica
EXPECTED	Rock	SVM		80			10	10	
		KNN	10	20		10			60
	Pop	SVM		50		10	10		30
		KNN	10			20		10	60
	Jazz	SVM		90			10		
		KNN	40	10		20			30
	Clásica	SVM		90					10
		KNN			10	20			10
	Salsa	SVM	20	70					10
		KNN	10			20			70
	Hip-hop	SVM	10	30	10		10	20	10
		KNN	30					1	3
	Electrónica	SVM	30	30	10	20		10	
		KNN	20	10	30	10	10		20

Imagen 10.7.1: Matriz de resultados prueba 7.

## 10.8 8° Prueba

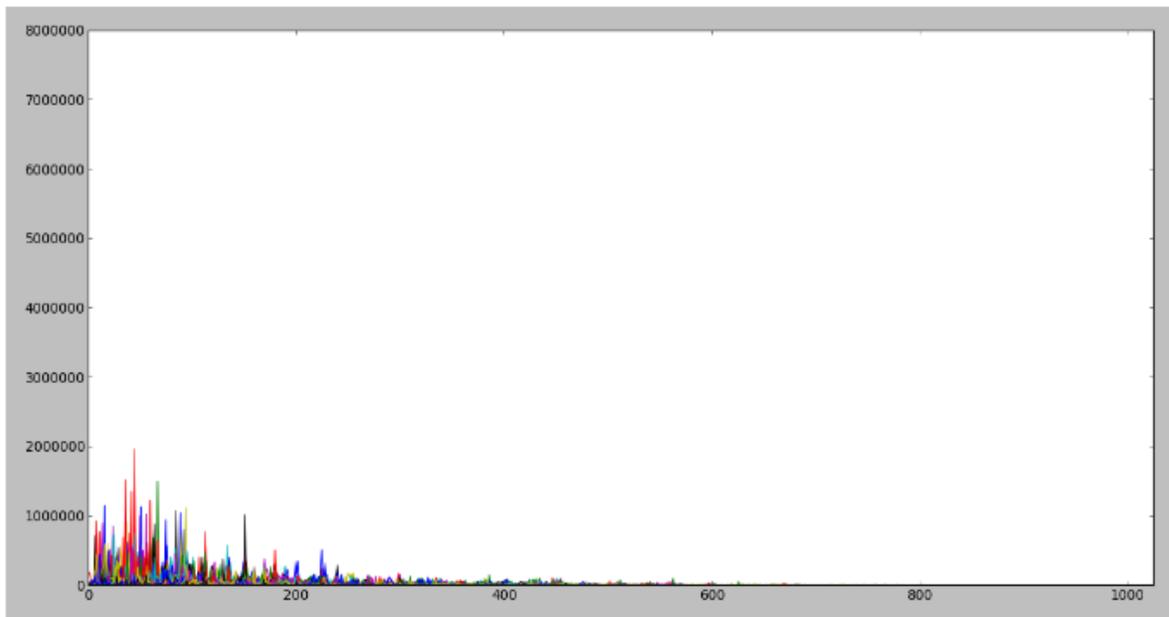
### 10.8.1 Contexto

Al analizar las pruebas anteriores se detectó que el vector de características que obtuvimos en el dominio del tiempo no son los apropiados para lograr clasificar las muestras por género musical, pues a simple vista no se puede observar algún patrón característico que contenga un grupo de canciones de un género en particular. Esta impresión 10.1 se confirmó al mento de hacer la primer prueba 10.2 con la que se puede deducir que para los clasificadores SVM y KNN en su espacio de clasificación no es posible las muestras de entrenamiento no se encuentran agrupadas. Para el espacio del KNN resultados 10.2 dejan ver que el género de clásica queda esparcido en todo el espacio haciendo que todas las predicciones sean que se pertenece al género clásica, algo similar sucede con el clasificador SVM pero con el género de hip-hop.

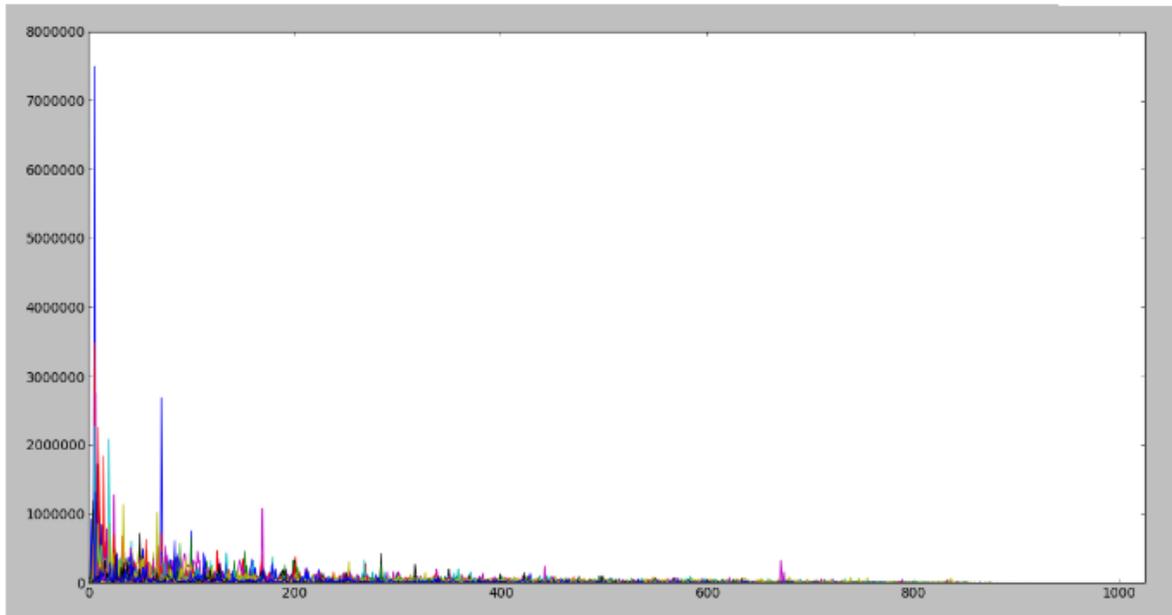
Por otra parte con las pruebas realizadas a partir de características en el dominio de la frecuencia los resultados obtenidos muestran mayor variación. Por este motivo se decide continuar desarrollando vectores de características en el dominio de la frecuencia e implementarlos en nuestros clasificadores.

## 10.8.2 Resultados

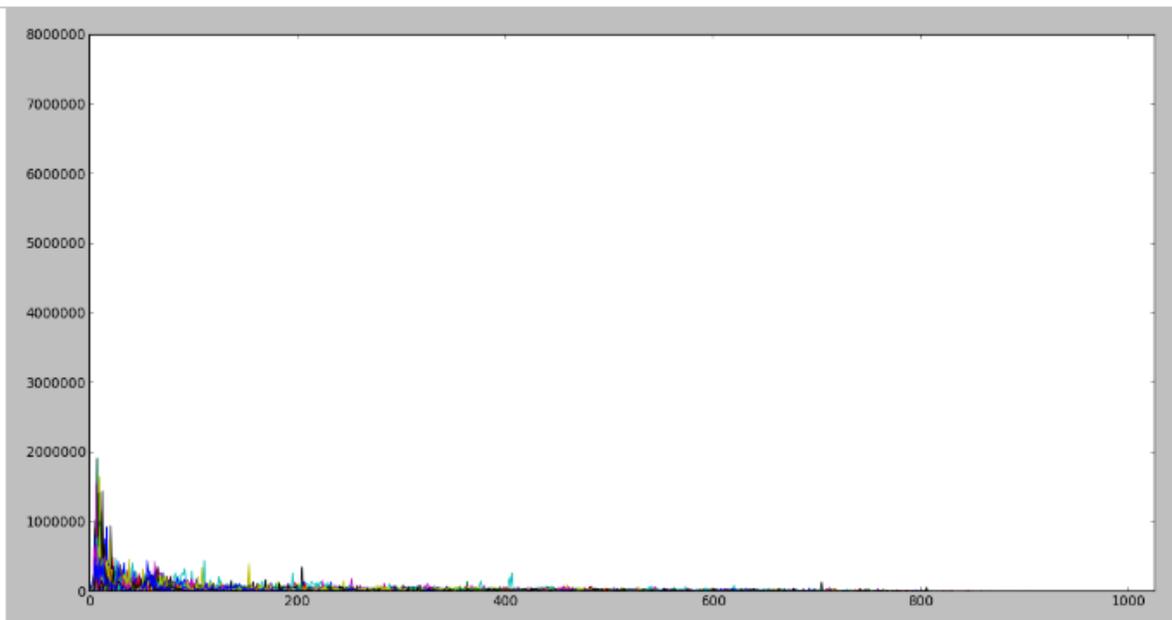
Las gráficas 10.8.1, 10.8.2, 10.8.3, 10.8.4, 10.8.5, 10.8.6, 10.8.7 muestran los conjuntos de entrenamiento de 50 canciones por género, por cada canción se obtienen 8,000 muestras que corresponden a los 5 segundos da la mitad de cada canción, de este segmento se realizan 39 ventanas de 2,048 muestras, posteriormente a cada ventana le aplicamos la FFT (se toman en cuenta las primeras 1024) y se promedian las 39 ventanas.



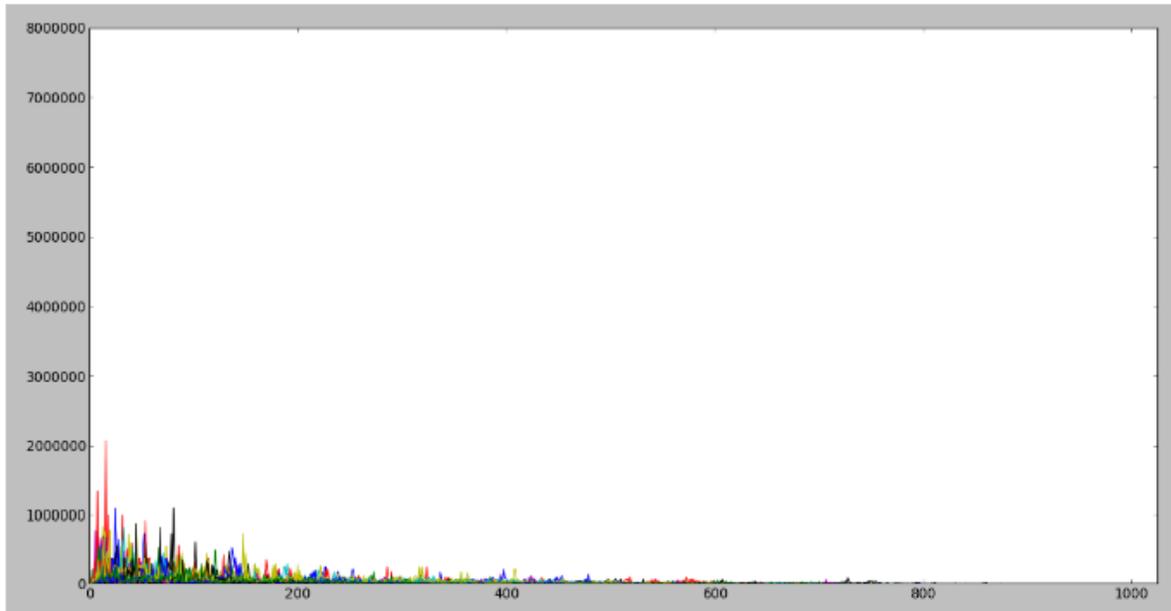
**Imagen 10.8.1:** Gráfica del género clásica.



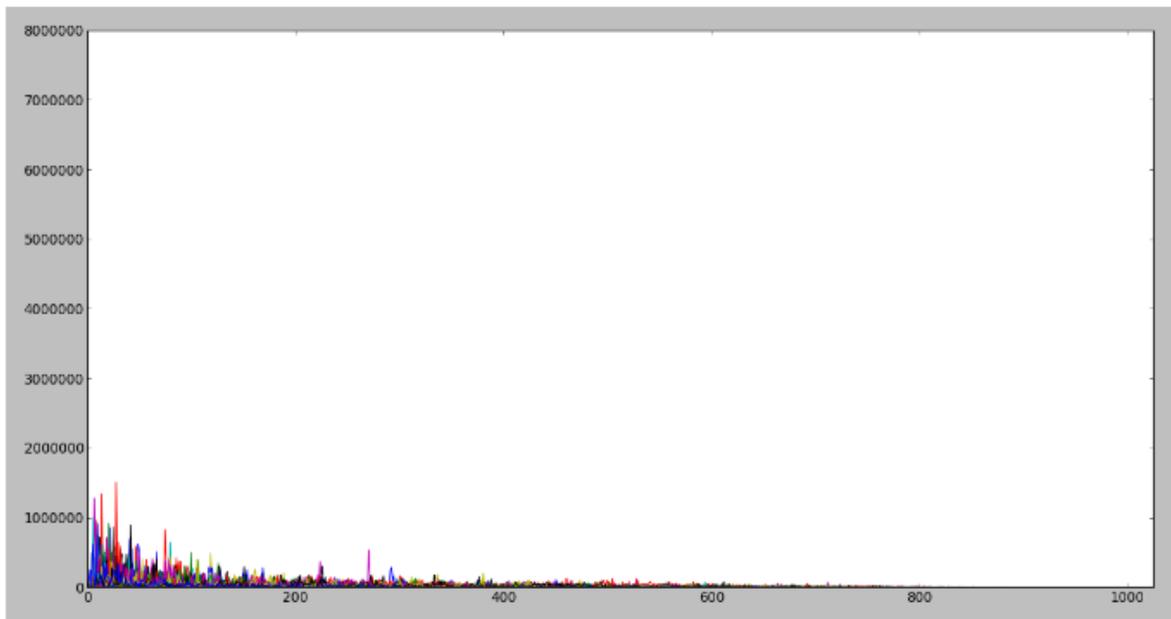
**Imagen 10.8.2:** Gráfica del género electrónica.



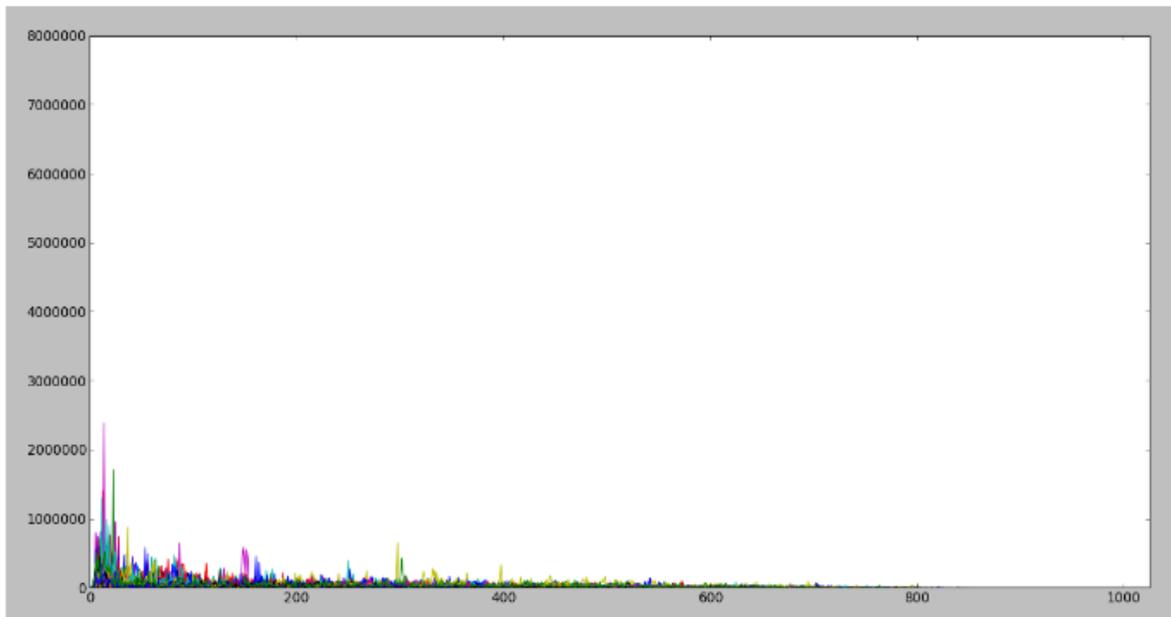
**Imagen 10.8.3:** Gráfica del género hip-hop.



**Imagen 10.8.4:** Gráfica del género jazz.



**Imagen 10.8.5:** Gráfica del género pop.

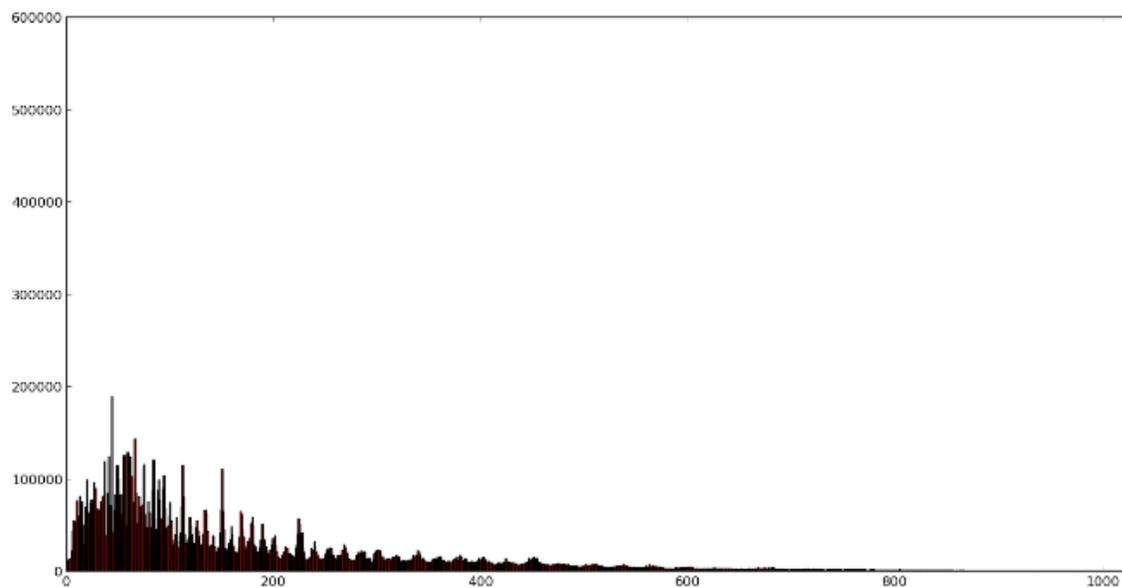


**Imagen 10.8.6:** Gráfica del género rock.

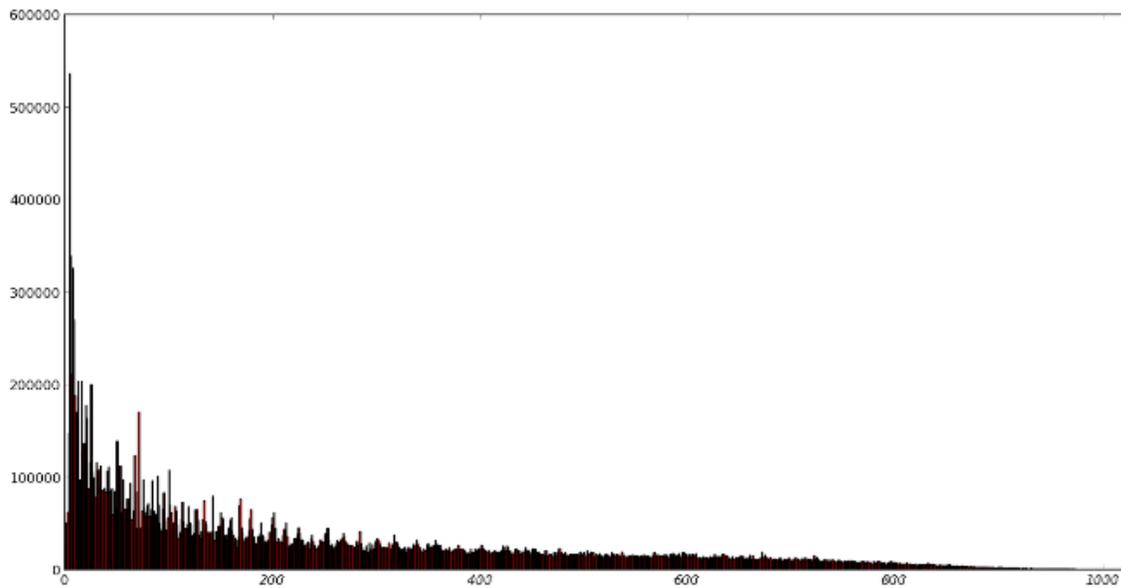


**Imagen 10.8.7:** Gráfica del género salsa.

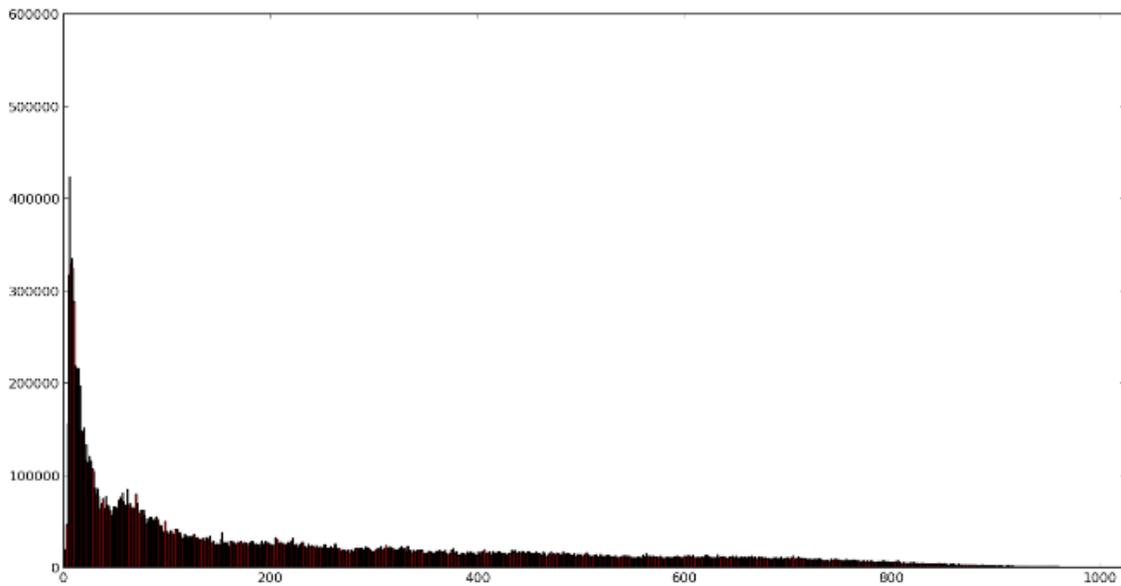
En las siguientes gráficas 10.8.8, 10.8.9, 10.8.10, 10.8.11, 10.8.12, 10.8.13, 10.8.14 se muestran los promedios por género de las muestras que se ilustran en las graficas anteriores (10.8.1, 10.8.2, 10.8.3, 10.8.4, 10.8.5, 10.8.6, 10.8.7).



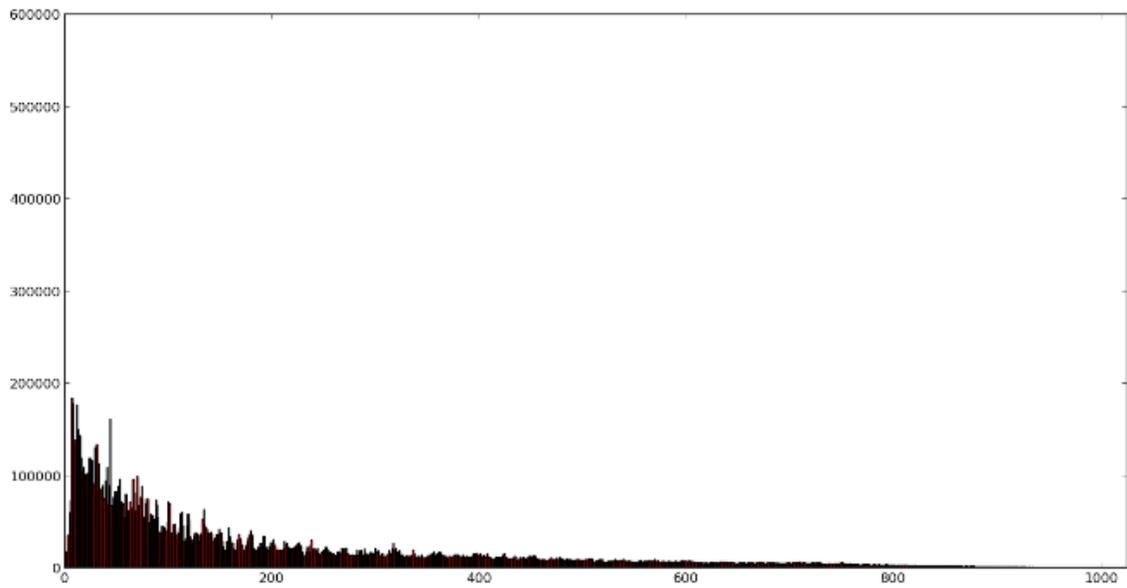
**Imagen 10.8.8:** Gráfica del promedio de muestras del género clásica.



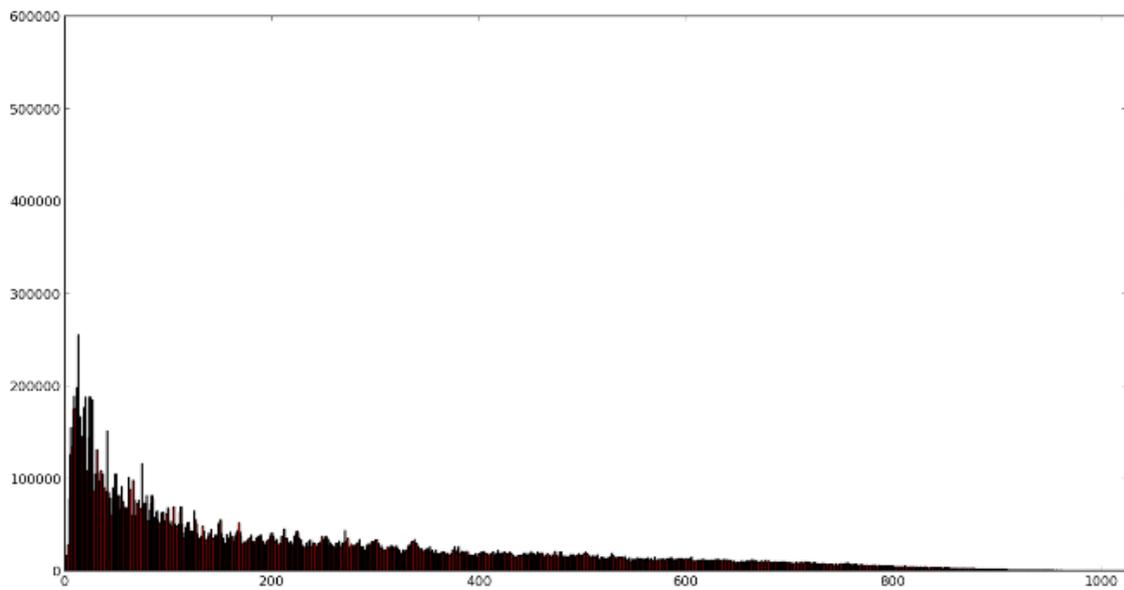
**Imagen 10.8.9:** Gráfica del promedio de muestras del género electrónica.



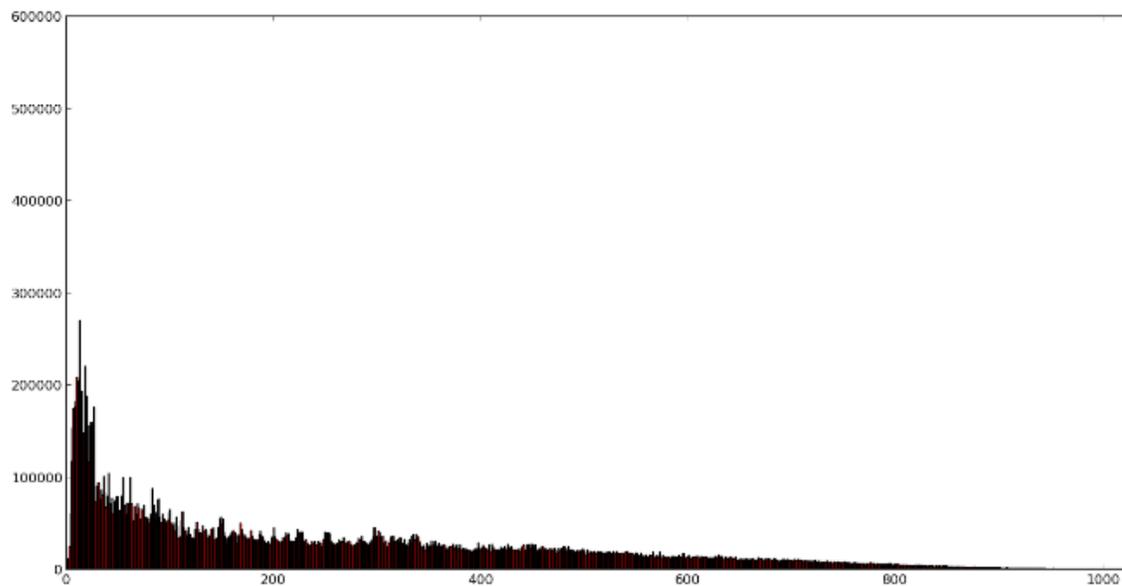
**Imagen 10.8.10:** Gráfica del promedio de muestras del género hip-hop.



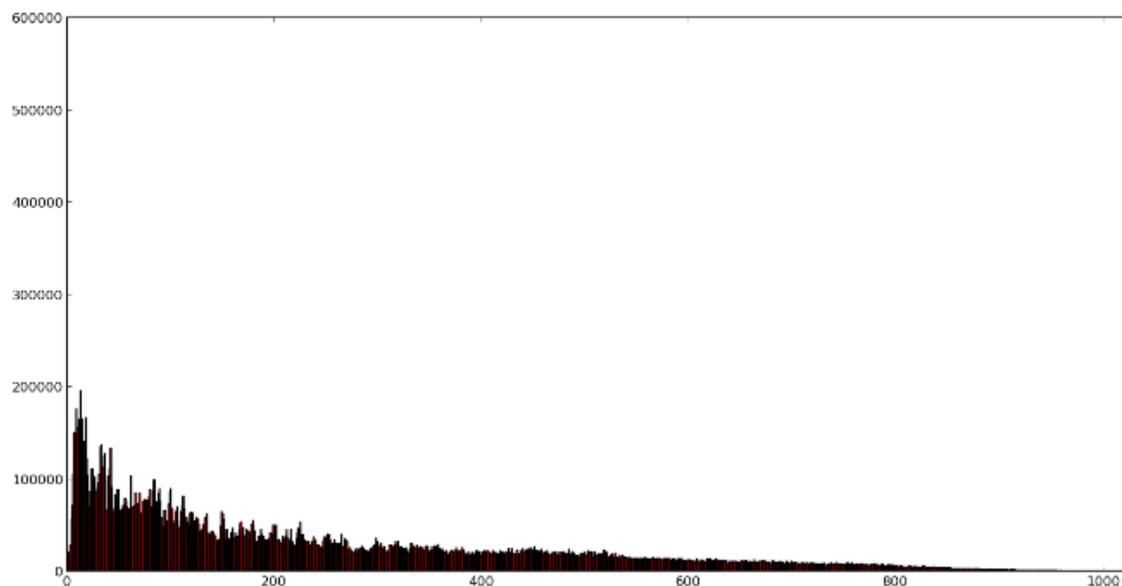
**Imagen 10.8.11:** Gráfica del promedio de muestras del género jazz.



**Imagen 10.8.12:** Gráfica del promedio de muestras del género pop.



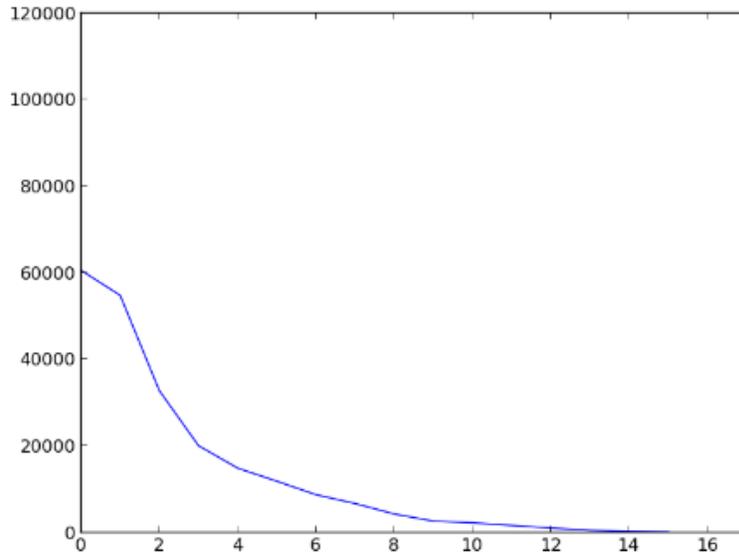
**Imagen 10.8.13:** Gráfica del promedio de muestras del género rock.



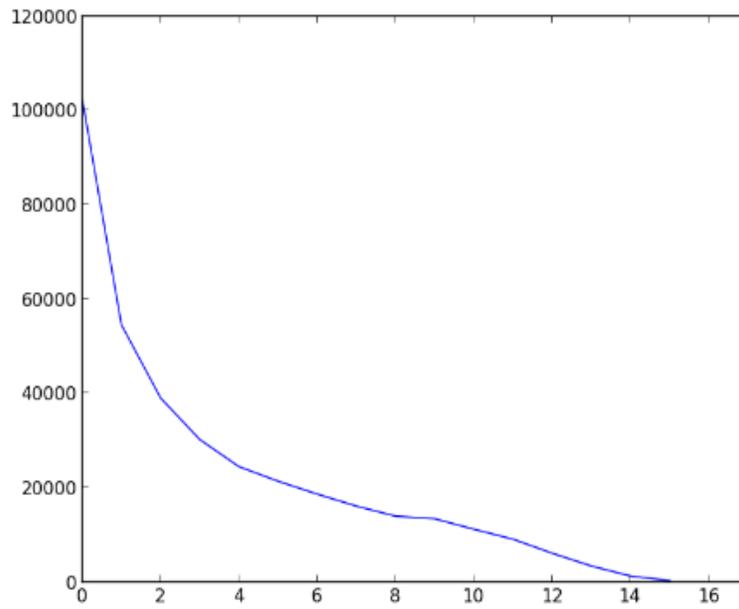
**Imagen 10.8.14:** Gráfica del promedio de muestras del género salsa.

De las muestras ilustradas anteriormente en las figuras 10.8.8, 10.8.9, 10.8.10, 10.8.11,

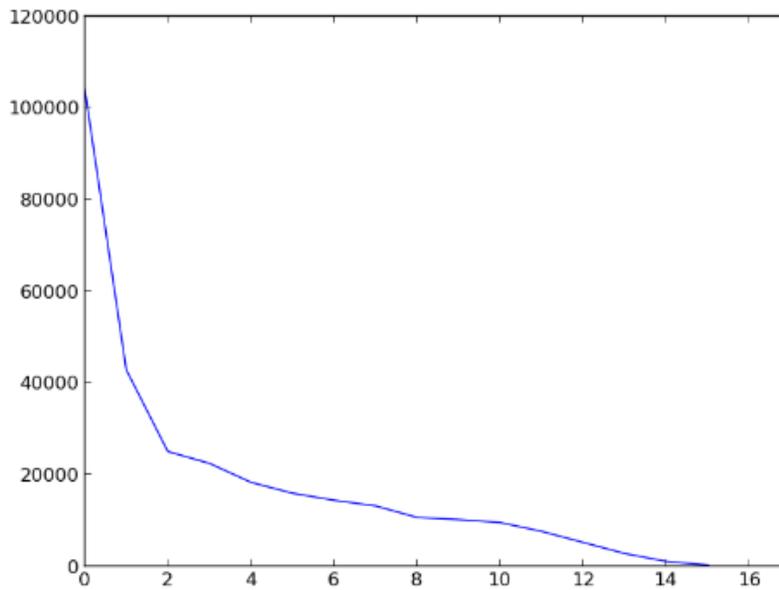
10.8.12, 10.8.13, 10.8.14, cada una de estas posee 1024 muestras, este número de características es reducido a 16 con el fin de lograr identificar con facilidad los rangos de la frecuencia más significativos en los géneros musicales. A continuación en las figuras 10.8.15, 10.8.16, 10.8.17, 10.8.18, 10.8.19, 10.8.20, 10.8.21 se grafica el resultado de esta reducción de número de características por cada género.



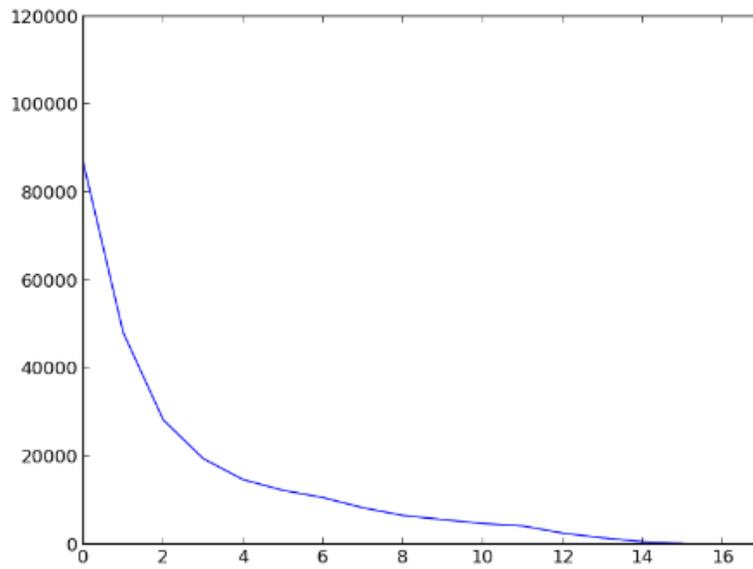
**Imagen 10.8.15:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género clásica.



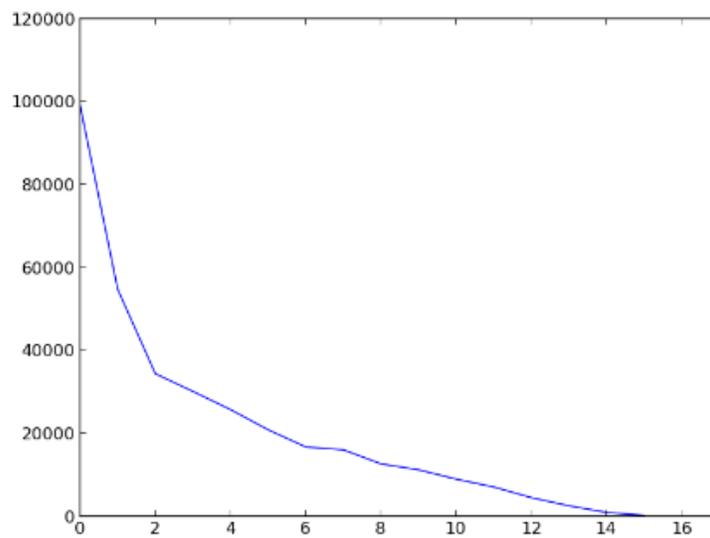
**Imagen 10.8.16:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género electrónica.



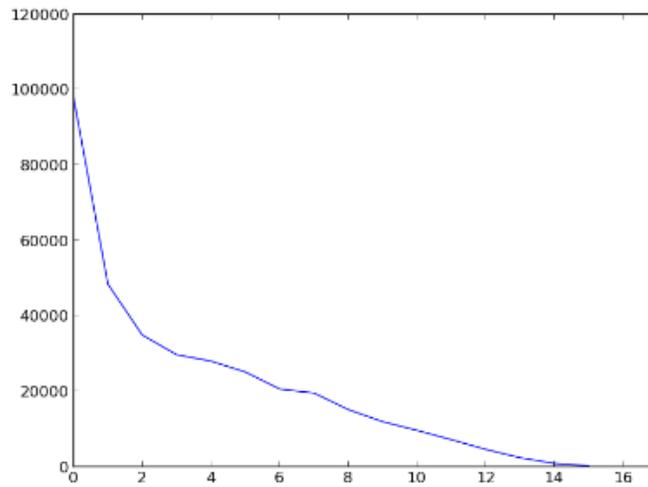
**Imagen 10.8.17:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género hip-hop.



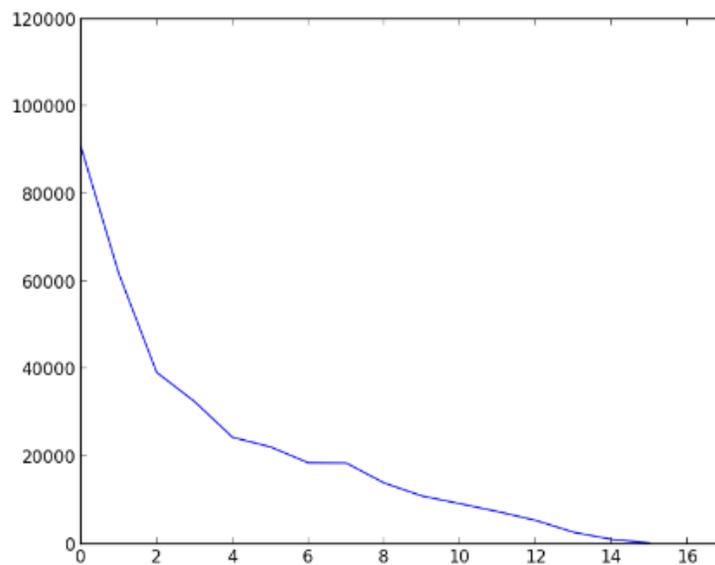
**Imagen 10.8.18:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género jazz.



**Imagen 10.8.19:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género pop.

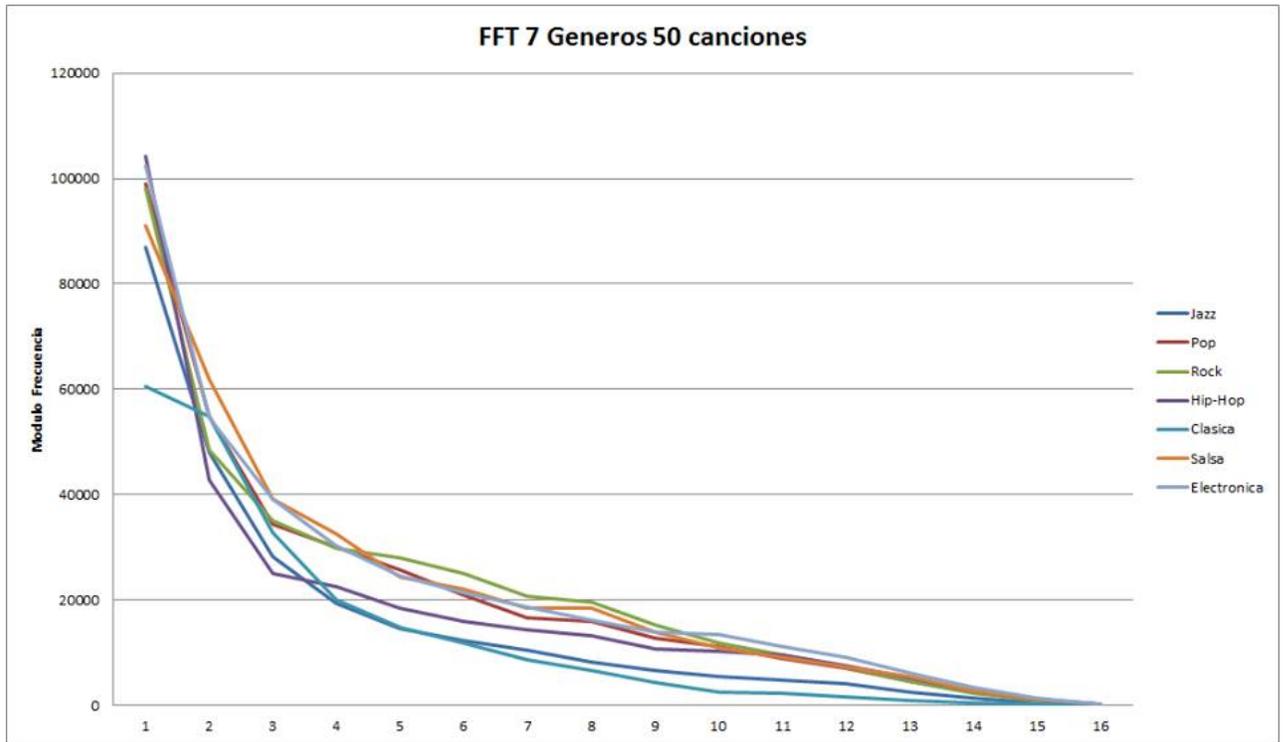


**Imagen 10.8.20:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género rock.



**Imagen 10.8.21:** Gráfica del promedio de muestras reducido de 1024 a 16 elementos del género salsa.

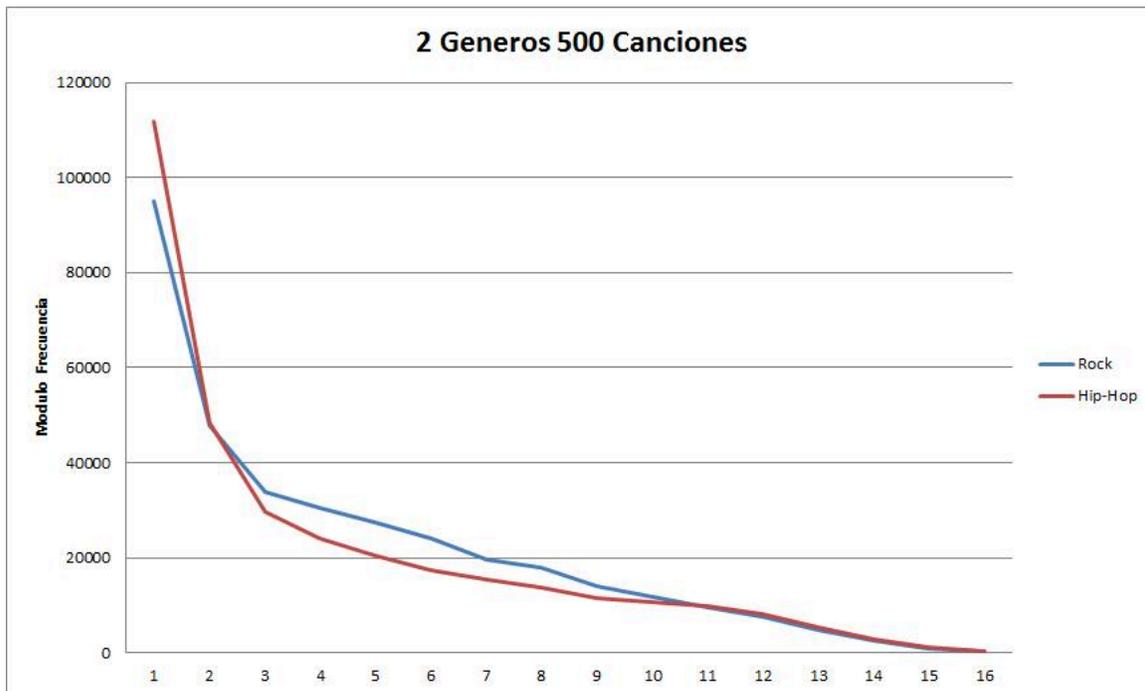
En la gráfica siguiente se puede observar la diferencia que existe de esta característica entre los géneros. 10.8.22



**Imagen 10.8.22:** Gráfica comparativa del promedio de muestras reducido de 1024 a 16 elementos de todos los géneros.

Al encontrar la diferencia entre estos géneros se decide continuar con pruebas enfocadas a la clasificación de estos. Como primer plan aumentaremos el número de muestras de entrenamiento para este género, estas pruebas son documentadas a continuación.

El grafico a continuación 10.8.23 deja ver con mayor claridad la diferencia que existe entre el género de rock y hip-hop.



**Imagen 10.8.23:** Gráfica comparativa del género rock y hip-hop.

## 10.9 9° Prueba

De acuerdo a la diferencia que existe entre los géneros rock y hip-hop se realizó un entrenamiento con 1000 muestras, 500 muestras correspondientes al género rock y 500 del género hiphop.

### 10.9.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

- ⇒ Prototipo I: 1.
- ⇒ Prototipo II: 6.
- ⇒ Prototipo III: 1 y 2.

## 10.9.2 Resultados

La gráfica 10.9.1 muestra los resultados obtenidos con 1024 características, en ella se puede observar que el clasificador SVM entregó un buen desempeño mientras que el KNN mostró un resultado muy pobre pues no tiene capacidad de decisión en la selección.

		Rock		Hip-hop	
		SVM	KNN	SVM	KNN
EXPECTED	Rock			82	100
	Hip-Hop	68	0		

**Imagen 10.9.1:** Gráfica de resultados contemplando 1024 características prueba 9.

Los resultados obtenidos con 2048 características se muestran en el gráfico 10.9.2 se puede observar tanto el clasificador SVM como el KNN mostraron resultados muy pobres con lo que nos damos cuenta de que este aumento en el vector de características confunde más a los clasificadores al momento de hacer la decisión de género.

		Rock		Hip-hop	
		SVM	KNN	SVM	KNN
EXPECTED	Rock			100	0
	Hip-hop	0	100		

**Imagen 10.9.2:** Gráfica de resultados contemplando 2048 características prueba 9.

## 10.10 10° Prueba

Como en la prueba anterior se realizó el entrenamiento con 500 muestras para el género de rock y 500 para el género de hiphop.

### 10.10.1 Contexto

Versiones de prototipos utilizadas para las pruebas:

⇒ Prototipo I: 1.

⇒ Prototipo II: 7.

⇒ Prototipo III: 1 y 2.

### 10.10.2 Resultados

Los resultados obtenidos con la variante de 8 muestras características del prototipo II versión 7 se muestran en la matriz 10.10.1.

		Rock		Hip-hop	
		SVM	KNN	SVM	KNN
EXPECTED	Rock			0	80
	Hip-hop	100	56		

**Imagen 10.10.1:** Gráfica de resultados contemplando 8 características prueba 10.

Los resultados obtenidos con la variante de 16 muestras características del prototipo II versión 7 se muestran en la matriz 10.10.2.

		Rock		Hip-hop	
		SVM	KNN	SVM	KNN
EXPECTED	Rock			88	88
	Hip-hop	50	50		

**Imagen 10.10.2:** Gráfica de resultados contemplando 16 características prueba 10.

## 10.11 11° Prueba

### 10.11.1 Contexto

Versiones utilizadas de prototipos.

⇒ Prototipo I: 1.

⇒ Prototipo II: 6.

⇒ Prototipo III: 1.

Para esta prueba se crearon árboles de decisión, en cada extremo se encuentra una pareja de géneros, de cada pareja se realiza una clasificación, la predicción resultante se vuelve a

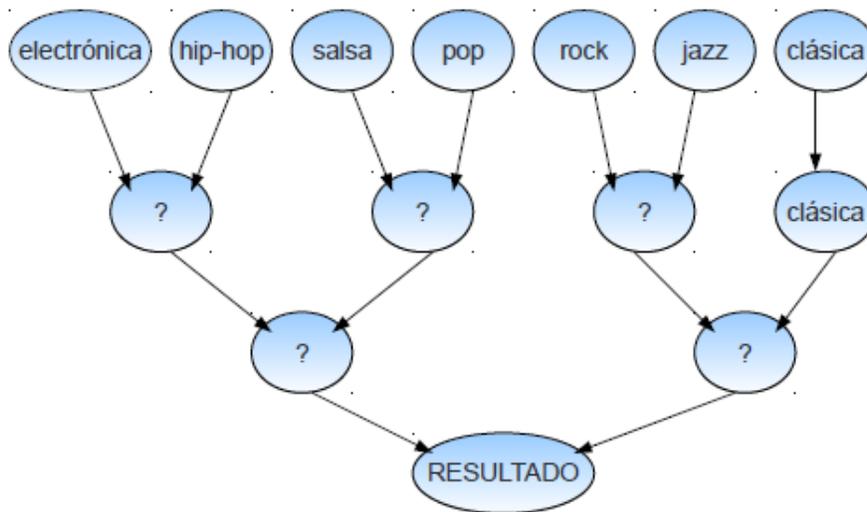
clasificar con el resultado de otro extremo.

El clasificador usado en cada par de géneros fue el KNN.

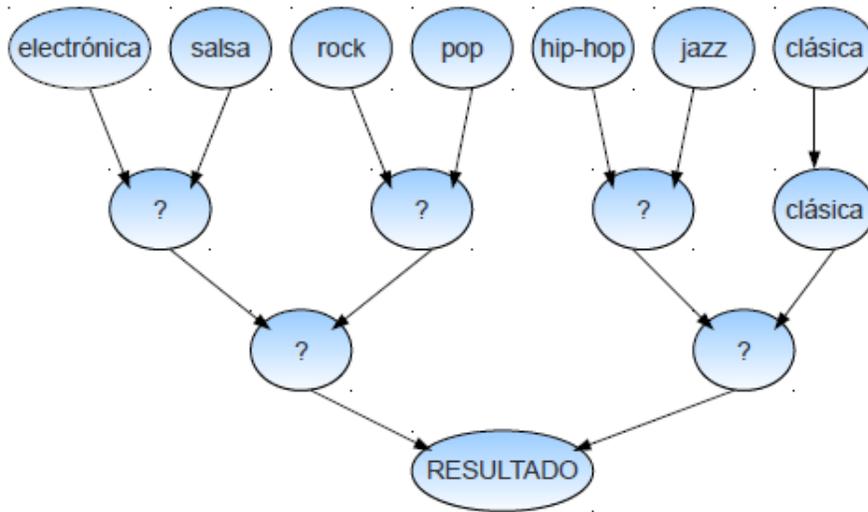
Se realizó un entrenamiento de 500 muestras por género.

El vector de características esta compuesto por 1024 elementos producto del promedio de las ventanas de FFT.

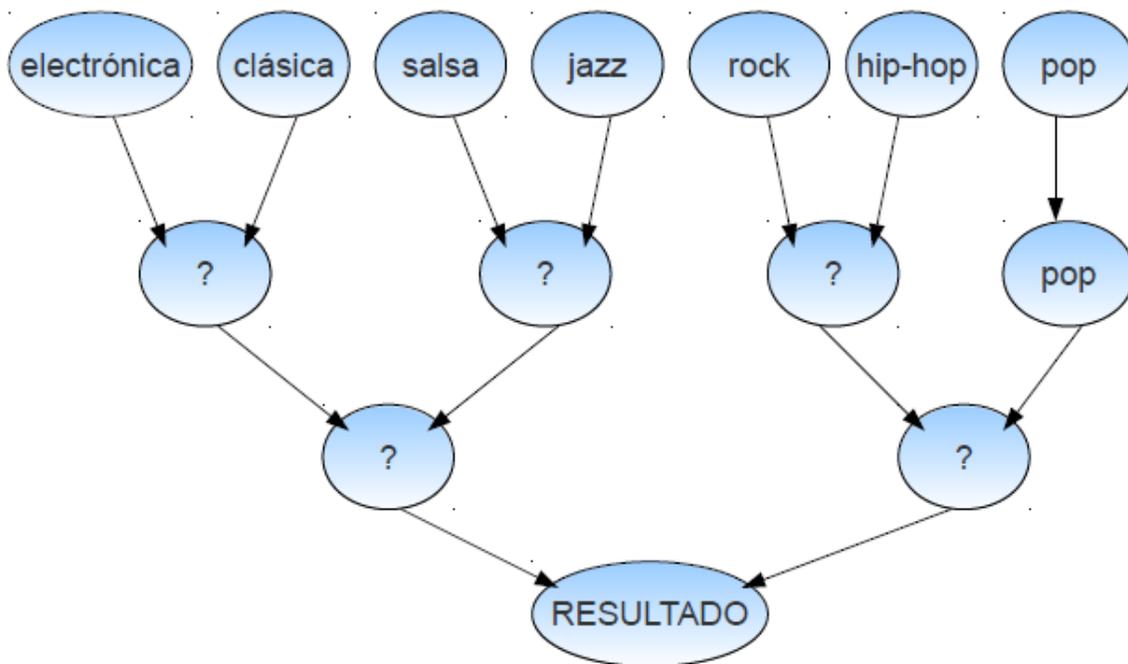
Los esquemas 10.11.1, 10.11.2 y 10.11.3 muestran las distintas configuraciones que se implementaron para el árbol de decisión.



**Imagen 10.11.1:** Primer árbol de decisión.



**Imagen 10.11.2:** Segundo árbol de decisión.



**Imagen 10.11.3:** Tercer árbol de decisión.

### 10.11.2 Resultados

Con cualquier árbol de decisión se obtienen los mismos resultados. Estos se muestran en la tabla de la figura 10.11.4.

Porcentajes		E	H	S	P	R	J	C	Total
C a n c i o n e s	Electronica	44	32	8	2	6	8	0	100
	Hip-Hop	28	38	2	12	14	4	2	100
	Salsa	0	8	58	10	4	18	2	100
	Pop	0	16	16	28	22	14	4	100
	Rock	4	10	12	10	50	8	6	100
	Jazz	2	4	20	6	6	56	6	100
	Clasica	0	0	0	2	2	50	46	100

**Imagen 10.11.4:** Gráfica de resultados realizando un árbol de decisión prueba 11.

## 10.12 12° Prueba

### 10.12.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 7.

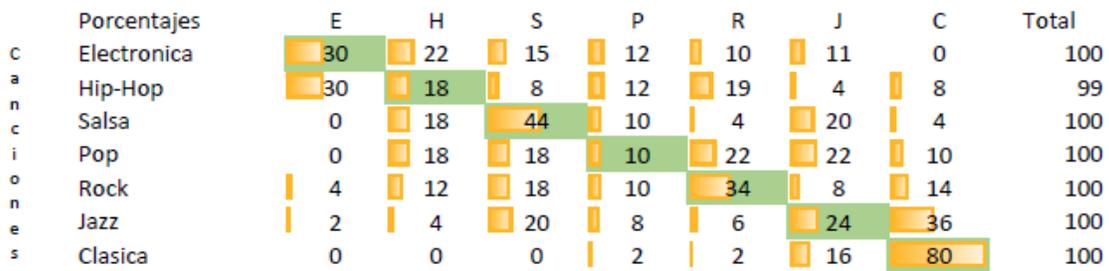
⇒ Prototipo III: 1 y 2.

### 10.12.2 Resultados

En esta prueba se utilizó la versión 7 del Prototipo II con su variante que arroja el vector de 16 características producto de la reducción de 1024 puntos de la FFT.

Ambos clasificadores (versión 1 y 2 del Prototipo III) dan los mismos resultados.

El entrenamiento se realizó con los 7 géneros con muestras de 500 elementos por cada uno de ellos. Los resultados obtenidos se muestran en la tabla del esquema 10.12.1.



**Imagen 10.12.1:** Gráfica de resultados contemplando 16 características.

## 10.13 13° Prueba

### 10.13.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 6.

⇒ Prototipo III: 1 y 2.

Para esta prueba se emplea la sexta versión del Prototipo II, que devuelve el vector de 1024 elementos producto del promedio de las FFT de las ventanas en que se fragmento el segmento de la pieza musical.

Se realizó el entrenamiento con los siete géneros con 500 muestras para cada uno. El prototipo II en su etapa de entrenamiento solo arroja una muestra por cada género de entrenamiento sin importar el número de muestras que participaron en el entrenamiento.

En esta prueba se obtienen también los mismos resultados con ambas versiones del prototipo III (clasificadores KNN o SVM).

### 10.13.2 Resultados

A continuación se muestran los resultados obtenidos en esta prueba en el gráfico de la figura

Porcentajes		E	H	S	P	R	J	C	Total
C a n c i o n e s	Electronica	42	32	8	2	6	8	0	100
	Hip-Hop	28	38	2	12	14	4	2	100
	Salsa	0	8	58	10	4	18	2	100
	Pop	0	16	16	26	22	16	4	100
	Rock	4	10	12	10	50	8	6	100
	Jazz	2	4	20	6	6	56	6	100
	Clasica	0	0	0	2	2	50	46	100

**Imagen 10.13.1:** Gráfica de resultados contemplando 1024 características y 500 muestras de entrenamiento por género.

## 10.14 14° Prueba

### 10.14.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 6.

⇒ Prototipo III: 1 y 2.

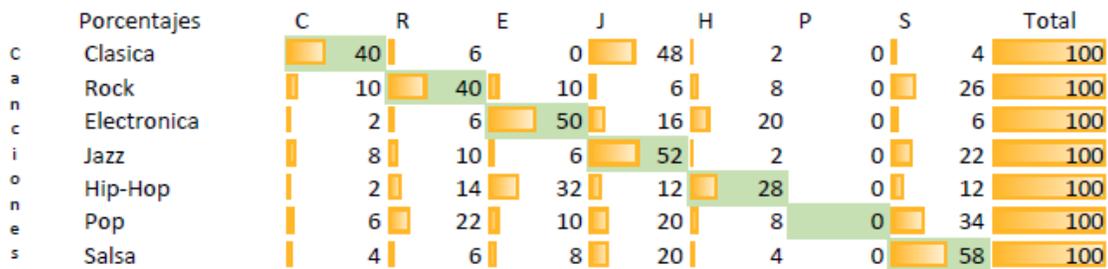
En esta prueba se hace uso de la 7 versión del Prototipo II que nos da un vector de 512 características, resultado del promedio de de las ventanas de un fragmento de la canción.

El entrenamiento se realiza con 500 muestras por cada género.

Como en las últimas pruebas los clasificadores KNN y SVM arrojan los mismos resultados.

### 10.14.2 Resultados

En la tabla de la figura 10.14.1 se muestran los resultados que se obtuvieron.



**Imagen 10.14.1:** Gráfica de resultados contemplando 512 características.

## 10.15 15° Prueba

### 10.15.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 9.

⇒ Prototipo III: 1 y 2.

El prototipo empleado en esta prueba toma 10 segundos de la mitad de la canción y crea ventanas de 1024 elementos para después calcular la FFT y se obtiene el módulo de cada elemento de la FFT. Se promedian todas los módulos de todas las ventanas y se obtiene un vector de características de 1024 elementos.

El entrenamiento se realiza con 500 muestras por cada género. Se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.

### 10.15.2 Resultados

En la tabla de la figura 10.15.1 se muestran los resultados que se obtuvieron.

	C	R	E	J	H	P	S
Clásica	44.2	0.6	0	51.3	0	0	3.9
Rock	5.27	50	2.78	10.7	7.5	12	11.75
Electrónica	0.55	6.7	53.75	15.25	16.5	0	7.25
Jazz	6.7	5.7	1.7	70.8	1.7	0	13.4
Hip-Hop	1.4	5.6	23.6	5.4	50.4	11.8	1.8
Pop	4.5	18.2	4.95	15.8	11.53	26.92	18.1
Salsa	3.1	8.2	2.3	17.3	8.4	10.4	50.3

**Imagen 10.15.1:** Gráfica de resultados contemplando el módulo de la FFT.

## 10.16 16° Prueba

### 10.16.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 9.

⇒ Prototipo III: 1 y 2.

El prototipo empleado en esta prueba toma 5 segundos de la mitad de la canción y crea ventanas de 1024 elementos para después calcular la FFT y se obtiene la fase de cada elemento de la FFT. Se promedian todas los módulos de todas las ventanas y se obtiene un vector de características de 1024 elementos.

El entrenamiento se realiza con 500 muestras por cada género. Se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.

### 10.16.2 Resultados

En la tabla de la figura 10.16.1 se muestran los resultados que se obtuvieron.

	C	R	E	J	H	P	S
Clásica	0	0	0	100	0	0	0
Rock	0	0	0	100	0	0	0
Electrónica	0	0	0	100	0	0	0
Jazz	0	0	0	100	0	0	0
Hip-Hop	0	0	0	100	0	0	0
Pop	0	0	0	100	0	0	0
Salsa	0	0	0	100	0	0	0

**Imagen 10.16.1:** Gráfica de resultados contemplando la Fase de la FFT.

## 10.17 17° Prueba

### 10.17.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 9.

⇒ Prototipo III: 1 y 2.

En esta prueba se emplea el vector de características de 20148 elementos que contiene la fase y el módulo de la FFT de ventanas de 1024 características.

El entrenamiento se realiza con 500 muestras por cada género. Se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.

### 10.17.2 Resultados

En la tabla de la figura 10.17.1 se muestran los resultados que se obtuvieron.

	C	R	E	J	H	P	S
Clásica	41.2	1.2	0	55.3	0	1.2	1
Rock	6.5	43.8	2.5	13.5	9.5	16.3	7.9
Electrónica	0.5	8.5	44.75	11.25	22	9.75	3.25
Jazz	7.3	7.1	0.33	63.76	1	9.7	10.7
Hip-Hop	1.4	7.4	27.2	2.2	44	14	3.8
Pop	6.3	21.5	5.4	17.4	12.2	24.9	12.2
Salsa	3.7	10.3	2.9	28.92	6	12.52	35.62

**Imagen 10.17.1:** Gráfica de resultados contemplando la fase y el módulo de la FFT.

## 10.18 18° Prueba

### 10.18.1 Contexto

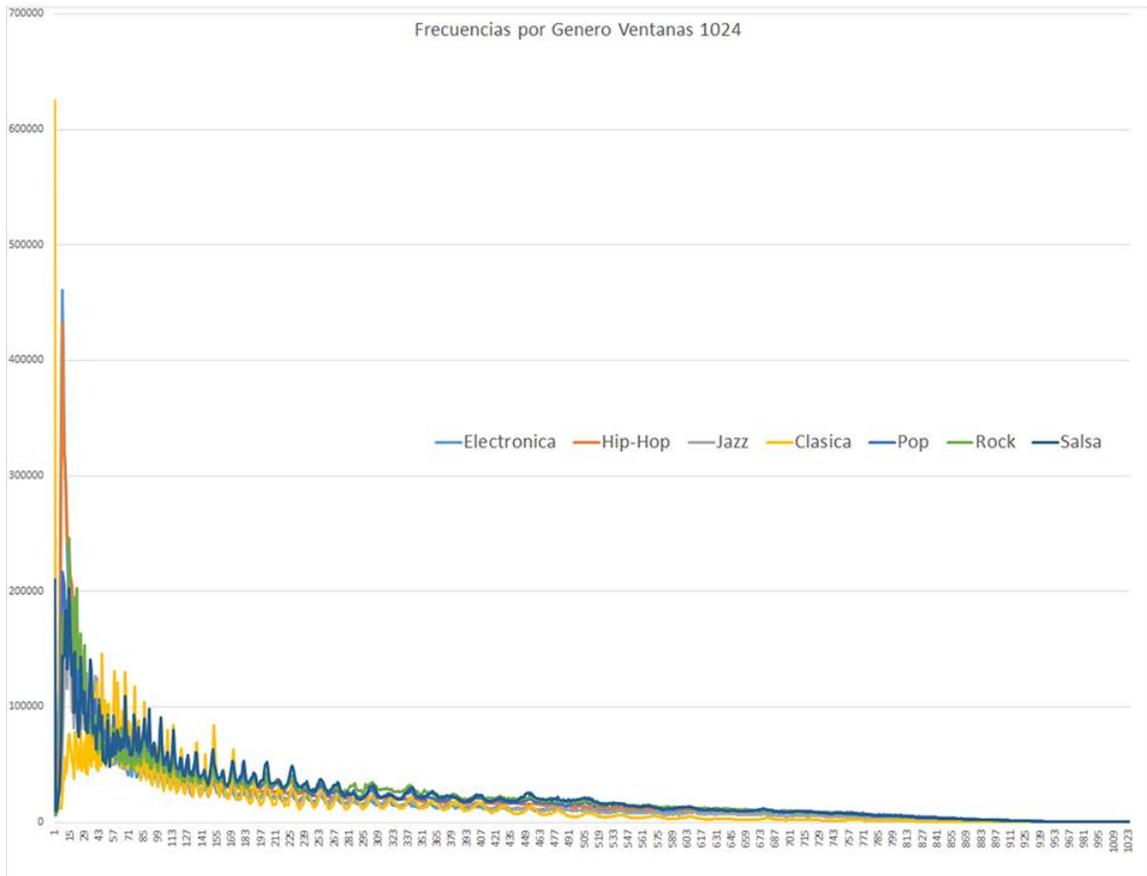
⇒ Prototipo I: 1.

⇒ Prototipo II: 8.

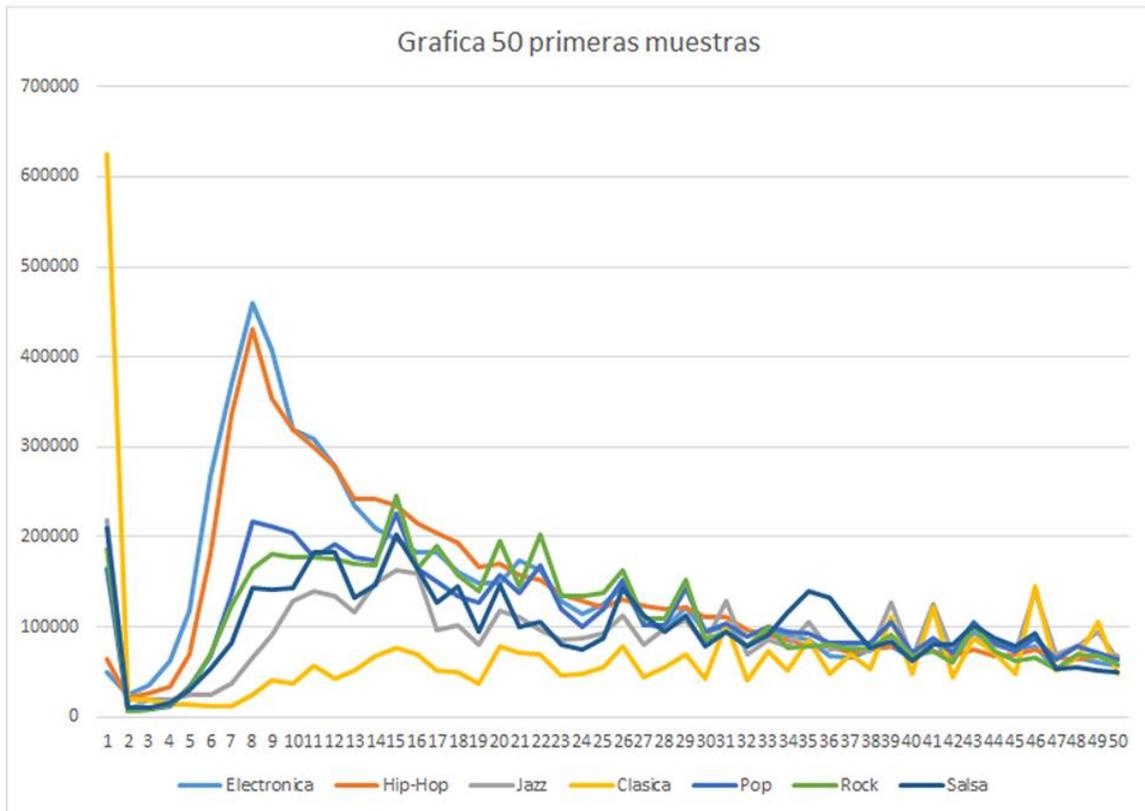
⇒ Prototipo III: 1 y 2.

De acuerdo a los resultados obtenidos en la prueba en la que se muestra una gráfica con los promedios de la FFT de los siete géneros, observamos que en los primeros 50 elementos es en donde existe una mayor diferencia. Por este motivo se decide realizar esta prueba para comprobar si realmente son rasgos que permitan una clasificación.

En la gráfica 10.18.1 se muestra la FFT promedio de 1024 elementos, mientras que en la gráfica 10.18.2 se muestran las primeras muestras que poseen mayor diferencia.



**Imagen 10.18.1:** Gráfica de FFT de 1024 puntos promedio por género.



**Imagen 10.18.2:** Primeros puntos con mayor diferencia de la FFT promedio de 1024 elementos.

El entrenamiento se realiza con 500 muestras por cada género. Se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.

### 10.18.2 Resultados

En la tabla de la figura 10.18.3 se muestran los resultados que se obtuvieron.

	C	R	E	J	H	P	S
Clásica	41.2	1.2	0	55.3	0	1.2	1
Rock	6.5	43.8	2.5	13.5	9.5	16.3	7.9
Electrónica	0.5	8.5	44.75	11.25	22	9.75	3.25
Jazz	7.3	7.1	0.33	63.76	1	9.7	10.7
Hip-Hop	1.4	7.4	27.2	2.2	44	14	3.8
Pop	6.3	21.5	5.4	17.4	12.2	24.9	12.2
Salsa	3.7	10.3	2.9	28.92	6	12.52	35.62

**Imagen 10.18.3:** Gráfica de resultados contemplando los primeros 50 elementos de la FFT.

## 10.19 19° Prueba

### 10.19.1 Contexto

⇒ Prototipo I: 1.

⇒ Prototipo II: 6.

⇒ Prototipo III: 1 y 2.

En la prueba que se obtuvieron mejores resultados se experimento con un vector de características de 1024 elementos producto del promedio de la FFT de las ventas de 2048 muestras de las que se toma solo la primer mitad 1024 muestras, en esta prueba se observan que los géneros clásica salsa y electrónica son los que presentan los puntajes de acertividad más altos y también se observa que hay géneros que causan mayor confusión al momento de intentar clasificar. Por este motivo en esta prueba solo se realiza el entrenamiento y las pruebas con los géneros clásica salsa y rock.

El entrenamiento se realiza con 500 muestras por cada género. Se promedian todos los elementos de los vectores de características de todas las muestras involucradas, obteniendo así una sola muestra de referencia por cada género entrenado.

## 10.19.2 Resultados

En la tabla de la figura 10.19.1 se muestran los resultados que se obtuvieron.

	Clasica	Electronica	Salsa
Clásica	67.87%	0	32.12%
Electrónica	2%	72.25%	25.75%
Salsa	3.60%	8.50%	87.82%

**Imagen 10.19.1:** Gráfica de resultados los tres mejores géneros.

# 11

## Referencias musicales

Resulta necesario tener un referente que indique o muestre que canciones pertenecen a qué género para obtener las muestras de entrenamiento y pruebas con las que se trabajan. En este capítulo se enlistan las referencias que se utilizan para la selección de muestras de cada género en la presente investigación.

### 11.1 Pistas de Entrenamiento

#### 11.1.1 Género clásica.

- ⇒ George Frideric Handel, Sinfonía: The Arrival of the Queen of Sheba: Solomon, HWV 67 (Op.8 No. 1)
- ⇒ Antonio Vivaldi, I. Allegro: The Four Seasons, Concerto No. 1 in ERV 269 'La primavera/Spring'
- ⇒ Johann Sebastian Bach, Toccata: Toccata and Fugue in D Minor, BWV 565

- ⇒ Wolfgang Amadeus Mozart, III. Rondo (Allegro vivace): Horn Concerto No. 4 in E flat K495
- ⇒ Ludwig Van Beethoven- Symphony No. 5 in C Minor, Op.67 - I. Allegro con brio
- ⇒ Pyotr Il'yich Tchaikovsky, 1812 Overture, Op.49 (conclusion)
- ⇒ Ludwig Van Beethoven, Ode to Joy from Symphony No. 9 in D minor Choral, Op.125
- ⇒ Edvard Grieg, I. Allegro molto moderato (excerpt): Piano Concerto in A Minor, Op.16
- ⇒ Jean Sibelius, III. Alla marcia: Karelia Suite, Op.11
- ⇒ Gustav Holst, Jupiter, the Bringer of Jolity (conclusion): The Planets, Op.32
- ⇒ Pyotr Il'yich Tchaikovsky, Piano Concerto No. 1 in B Flat minor, Op.23 - I. Allegro non troppo e molto maestoso (opening)
- ⇒ Jean Sibelius, Finlandia, Symphonic Poem, Op.26
- ⇒ Giuseppe Verdi, Triumphal Chorus & Grand March (Gloria all Egitto): Aida
- ⇒ Sergey Prokofiev, Montagues and Capulets (Dance of the Knights)
- ⇒ Johann Strauss I, Radetzky March, Op.228
- ⇒ Edward Elgar, Pomp and Circumstance March No. 1 in D (Land of Hope and Glory) (conclusion)
- ⇒ Sergei Rachmaninov, I. Allegro ma non tanto (excerpt): Piano Concerto No.3 in D minor
- ⇒ Camille Saint-Saens, Maestoso Allegro: Symphony No. 3, 'Organ' Op. 78
- ⇒ Sergei Rachmaninov, II. Adagio sostenuto (excerpt): Piano Concerto No.2 in C Minor, Op.18
- ⇒ Edvard Grieg, Prelude (Morning): Peer Gynt
- ⇒ Dmitri Shostakovich, Romance from the Gadfly
- ⇒ Samuel Barber, Adagio for Strings, Op.11
- ⇒ Nikolay Andreyevich Rimsky-Korsakov, The Story of the Kalender Prince (opening): Scheherazade, Op.35
- ⇒ Jules Massenet, Méditation: Thaïs
- ⇒ Alexander Porfir'yevich Borodin, Polovtsian Dances (beginning) from Prince Igor
- ⇒ Aram Khachaturian, Adagio: Spartacus Ballet Suite No. 2
- ⇒ Pyotr Il'yich Tchaikovsky, Fantasy Overture - Romeo and Juliet (extract)
- ⇒ Ludwig Van Beethoven, II. Allegretto (abbrev. version) from Symphony No. 7 in A Op. 92
- ⇒ Johann Sebastian Bach, Largo ma non tanto: Concerto for 2 Violins in D Major, BWV 1043
- ⇒ John Williams, Schindler's List
- ⇒ Pyotr Il'yich Tchaikovsky, V. Danse des mirlitons (Dance of the Reed Pipes): The Nutcracker - Ballet, Op.71 No. 12 Divertissement (Act II)
- ⇒ Ludwig Van Beethoven, V. Allegretto (Shepherds' Hymn): Symphony No. 6 in F, Op.68 'Pastoral'
- ⇒ George Frideric Handel, Zadok the Priest
- ⇒ Johann Sebastian Bach, Zion hört die Wächter singen (Cantata No. 140)

- ⇒ Gregorio Allegri, Miserere a 9
- ⇒ Franz Schubert, Zum Sanctus: Deutsche Messe D. 872 (mit "Gebet des Herrn")
- ⇒ Cesar Franck, Panis angelicus
- ⇒ Hector Berlioz, L'Enfance du Christ Op. 25: L'Adieu des bergers (Shepherds' Farewell)
- ⇒ Gabriel Faure, Requiem, Op.48: In Paradisum
- ⇒ Charles Gounod, Sanctus: St Cecilia Mass
- ⇒ Wolfgang Amadeus Mozart, Laudate Dominum: Vesperae solennes de confessore (Solemn Vespers), K.339
- ⇒ John Milford Rutter, Cantique De Jean Racine

- ⇒ Giuseppe Verdi, Ingemisco: Messa da Requiem, Sequenza:
- ⇒ Wolfgang Amadeus Mozart, Ave verum corpus, K.618
- ⇒ George Frideric Handel, I know that my Redeemer liveth (soprano): Messiah highlights
- ⇒ Joseph Haydn, The heavens are telling (The Creation: Part 1)
- ⇒ Gabriel Faure, IV. Pie Jesu: Requiem, Op.48
- ⇒ Wolfgang Amadeus Mozart, 6. Lacrimosa: Requiem in D Minor, K.626, III. Sequenz
- ⇒ George Frideric Handel, Hallelujah (Chorus): Messiah highlights

2

### 11.1.2 Género electrónica.

- ⇒ Ten - Sander Van Doorn, Mark knight & Underworld
- ⇒ Rage - Deniz Koyu
- ⇒ Symphonica - Nicky Romero
- ⇒ Here & now - Ummet Ozcan
- ⇒ Chasing Summers -Tiesto
- ⇒ Redux - DubVision
- ⇒ Sultan & Ned Shepard & Fedde Le Grand - Long Way From Home (Original Mix)
- ⇒ Cedric Gervais & Borgor - Deception (Original Mix)
- ⇒ David Guetta Feat. Ne-Yo & Akon - Play Hard (Albert Neve Remix)
- ⇒ Hard Rock Sofa - Rasputin (Original Mix)
- ⇒ Firebeatz - Gangster (Original Mix)
- ⇒ Moguai - Champs (Original Mix)
- ⇒ Madeon Feat. Ellie Goulding - Stay Awake (Original Mix)

- ⇒ Moguai & Dimitri Vegas & Like Mike - Mammoth (Original Mix)
- ⇒ Bassjackers & Kenneth G - Duckface (Original Mix)
- ⇒ Vangelis, Blade Runner (End Titles) (1982)
- ⇒ New Order, Bizarre Love Triangle (1986)
- ⇒ Talk Talk, Living In Another World (1986)
- ⇒ Massive Attack, Unfinished Sympathy (1991)
- ⇒ Decadance, On And On (Fears Keep On) (1983)
- ⇒ Hyetal, Beach Scene (2011)
- ⇒ Electrosexual & Mz Sunday Luv, I Feel Love
- ⇒ M83, Couleurs (2008)
- ⇒ Pet Shop Boys, One More Chance (1987)
- ⇒ Björk, Hyperballad (1995)
- ⇒ Don't you worry child
- ⇒ Ill listen, Armin Van Buuren
- ⇒ J ai envie de toi,
- ⇒ Greyhound, Swedish House
- ⇒ In my mind (Axwel Remix), David Guetta
- ⇒ Titatnium (Alesso Remix), David Guetta
- ⇒ Pair of dice, Tiesto, Allure
- ⇒ Levels (SKRILLEX REMIX), Avicci
- ⇒ J ai envie de toi (Protoculture Remix), Armin Van Buuren
- ⇒ Perfect Love, Aly & Fila
- ⇒ The ocean, Paul Van Dyk
- ⇒ World falls apart, Dash Berlin, Jonathan Mendelsohn
- ⇒ First of the year (EQUINOX), Skrillex.
- ⇒ City of dreams, Dirty South, Alesso
- ⇒ Somebody that I used to know, Gotye, Kimbra
- ⇒ Ten, Sander Van Doorn, Mark knight & Underworld
- ⇒ Rage, Deniz Koyu
- ⇒ Symphonica, Nicky Romero
- ⇒ Here & now, Ummet Ozcan
- ⇒ Chasing Summers, Tiesto
- ⇒ Redux, DubVision
- ⇒ Sultan & Ned Shepard & Fedde Le Grand, Long Way From Home (Original Mix)
- ⇒ Cedric Gervais & Borgor, Deception (Original Mix)
- ⇒ David Guetta Feat. Ne-Yo & Akon, Play Hard (Albert Neve Remix)
- ⇒ Hard Rock Sofa, Rasputin (Original Mix)
- ⇒ Firebeatz, Gangster (Original Mix)
- ⇒ Moguai, Champs (Original Mix)
- ⇒ Madeon Feat. Ellie Goulding, Stay Awake (Original Mix)
- ⇒ Moguai & Dimitri Vegas & Like Mike, Mammoth (Original Mix)
- ⇒ Bassjackers & Kenneth G, Duckface (Original Mix) item Pietro Mascagni, Intermezzo: Cavalleria Rusticana

### 11.1.3 Género hip-hop.

- ⇒ Geto Boys, Mind Playin' Tricks
- ⇒ Wu-Tang Clan, C.R.E.A.M.
- ⇒ Scarface, I Seen A Man Die
- ⇒ Notorious B.I.G., Juicy
- ⇒ 2Pac, Old School
- ⇒ Jay-Z, Dead Presidents II
- ⇒ Luniz, I Got 5 On It
- ⇒ Method Man & Redman, Da Rockwilder
- ⇒ Notorious B.I.G., Everyday Struggle
- ⇒ Pete Rock and CL Smooth, T.R.O.Y. (They Reminisce Over You)
- ⇒ Nas, The Message
- ⇒ 2Pac, Krazy
- ⇒ Pharcyde, Back In The Day
- ⇒ Snoop Dogg, Gin N' Juice
- ⇒ Notorious B.I.G., I Got a Story to Tell
- ⇒ 2Pac, Pour Out A Little Liquor
- ⇒ Souls of Mischief, 93 'Til Infinity
- ⇒ Black Moon, I Gotcha' Opin (Remix)
- ⇒ Eric B. & Rakim, Paid In Full
- ⇒ Mobb Deep, Shook Ones Pt. II
- ⇒ Bone Thugs-N-Harmony, Thuggish Ruggish Bone
- ⇒ Nas, Memory Lane
- ⇒ Gang Starr, Mass Appeal
- ⇒ Dead Prez, It's Bigger Than Hip Hop
- ⇒ Notorious B.I.G., Things Done Changed
- ⇒ 2Pac, So Many Tears
- ⇒ Slick Rick, Children's Story
- ⇒ Big Daddy Kane, Ain't No Half Steppin
- ⇒ Snoop Dogg, Nuthin But a G Thang ft. Dr. Dre
- ⇒ A Tribe Called Quest, Can I Kick It
- ⇒ Bone Thugs-N-Harmony, Crossroads
- ⇒ Notorious B.I.G., The What ft. Method Man
- ⇒ Warren G, G Funk Era
- ⇒ KRS One, MC's Act Like They Don't Know
- ⇒ Nas, NY State of Mind
- ⇒ Eminem, Lose Yourself
- ⇒ Pharoahe Monch, The Truth ft. Common & Talib Kweli
- ⇒ Crooklyn Dodgers '95, Return of the Crooklyn Dodgers
- ⇒ Blackstar, Mos Def & Talib Kweli, Thieves In The Night
- ⇒ Outkast, Players Ball
- ⇒ 2Pac, Keep Ya Head Up
- ⇒ De La Soul, Ego Trippin
- ⇒ GZA Genius, Shadowboxin' ft. Method Man
- ⇒ Common, The 6th Sense

- ⇒ Black Sheep, The Choice is Yours
- ⇒ Notorious B.I.G., Notorious Thugz ft. Bone Thugs-N-Harmony

- ⇒ DMX, Slippin
- ⇒ 2Pac, I Get Around
- ⇒ Naughty By Nature, Hip Hop Hooray
- ⇒ Big L, Put It On

#### 11.1.4 Género jazz.

- ⇒ Strange Fruit, (Billie Holliday, Cassandra Wilson)
- ⇒ What a Wonderful World, (Louis Armstrong, Lionel Hampton)
- ⇒ Round' Midnight, (Thelonious Monk, Miles Davis)
- ⇒ Goodbye Pork Pie Hat, (Charles Mingus, Stanley Clarke)
- ⇒ Mood Indigo, (Duke Ellington, Charles Mingus)
- ⇒ Naima, (John Coltrane, Woody Herman)
- ⇒ Body and Soul, (Coleman Hawkins, Dexter Gordon)
- ⇒ Willow Weep for Me, (Dexter Gordon, Clifford Brown)
- ⇒ I Remember Clifford, (Lee Morgan, Dizzy Gillespie)
- ⇒ Summertime, (Sidney Bichet, Herbie Hancock/Joni Mitchell)
- ⇒ Unforgettable, (Nat King Cole, Natalie Cole)
- ⇒ Black and Blue, (Louis Armstrong, Dinah Washington)
- ⇒ Time After Time, (Jay Jay Johnson, Sonny Stitt)
- ⇒ Lush Life, (Joe Henderson/Billy Strayhorn, Chick Corea)
- ⇒ Darn that Dream, (Dexter Gordon, Billie Holliday)
- ⇒ Im Old Fashioned, (John Coltrane, Ella Fitzgerald)
- ⇒ Moon River, (Louis Armstrong, Art Blakey)
- ⇒ Autumn in New York, (Billie Holliday, Charlie Parker)
- ⇒ Blue In Green, (Miles Davis/Bill Evans, Charlie Haden)
- ⇒ Mona Lisa, (Nat King Cole, Bing Crosby)
- ⇒ Dinah Washington, Mad About the Boy
- ⇒ Duke Ellington, Take The "A" Train
- ⇒ Stan Getz, Desafinado
- ⇒ The Oscar Peterson Trio, C Jam Blues
- ⇒ Sarah Vaughan, Lullaby Of Birdland
- ⇒ Jimmy Smith, Walk On The Wild Side
- ⇒ George Benson, Footin' It
- ⇒ Count Basie, Help
- ⇒ Sonny Rollins, Way out West

- ⇒ Louis Armstrong And The All-Stars, Basin Street Blues
- ⇒ Coleman Hawkins, Body And Soul
- ⇒ Nina Simone, See-Line Woman
- ⇒ Herbie Hancock, Watermelon Man
- ⇒ Ella Fitzgerald, Summertime
- ⇒ Charles Mingus, Mood Indigo
- ⇒ Bud Powell, Parisian Thoroughfare
- ⇒ Billie Holiday, Lady Sings The Blues
- ⇒ Bill Evans, Waltz For Debby
- ⇒ Let's do it, Ella Fitzgerald

- ⇒ I got Plenty o' Nuttin, Ella Fitzgerald
- ⇒ Embraceable you, Sarah Vaughan
- ⇒ It don't mean a thing, Carmen McRae
- ⇒ You and the night and the music, Helen Merrill
- ⇒ Don't explain, Nina Simone
- ⇒ Stormy Blues, Billie Holiday
- ⇒ It could happen to you, Chet Baker
- ⇒ My funny Valentine, Shirley Horn
- ⇒ Lonely Woman, Dee Dee Bridgewater
- ⇒ The look of love, Diana Krall
- ⇒ Africana, Dizzy Gillespie

#### 11.1.5 Género pop.

- ⇒ Beatles, Yesterday
- ⇒ Rolling Stones, Satisfaction
- ⇒ Nirvana, Smells Like Teen Spirit
- ⇒ Madonna, Like a Virgin
- ⇒ Michael Jackson, Billie Jean
- ⇒ Beatles, I Want To Hold Your Hand
- ⇒ Aretha Franklin, Respect
- ⇒ U2, With or Without You
- ⇒ Jackson 5, I Want You Back
- ⇒ Backstreet Boys, I Want It That Way
- ⇒ The Eagles, Hotel California

- ⇒ The Supremes, Where Did Our Love Go
- ⇒ Guns 'N Roses, Sweet Child O' Mine
- ⇒ Rolling Stones, Brown Sugar
- ⇒ John Lennon, Imagine
- ⇒ Sinead O'Conner, Nothing Compares 2 U
- ⇒ Stevie Wonder, Superstition
- ⇒ REM, Losing My Religion
- ⇒ Madonna, Vogue
- ⇒ Bob Dylan, Like A Rolling Stone
- ⇒ Van Morrison, Brown Eyed Girl
- ⇒ Michael Jackson, Beat It

- ⇒ Roy Orbison, Oh, Pretty Woman
- ⇒ Marvin Gaye, What's Going On
- ⇒ Britney Spears, Baby One More Time
- ⇒ Fleetwood Mac, Go Your Own Way
- ⇒ Prince, When Doves Cry
- ⇒ Hanson, Mmmhlop
- ⇒ Queen, Bohemian Rhapsody
- ⇒ Elton John, Your Song
- ⇒ Santana, Smooth
- ⇒ Otis Redding, Sitting On The Dock Of The Bay
- ⇒ The Who, My Generation
- ⇒ Alanis Morissette, Ironic
- ⇒ Bruce Springsteen, Born To Run
- ⇒ Tlc, Waterfalls

- ⇒ Naughty By Nature, OPP
- ⇒ David Bowie, Changes
- ⇒ Googoo Dolls, Iris
- ⇒ Whitney Houston, I Will Always Love You
- ⇒ Creedence Clearwater Revival, Proud Mary
- ⇒ Police, Every Breath You Take
- ⇒ Rolling Stones, Miss You
- ⇒ Abba, Dancing Queen
- ⇒ Eric Clapton, Tears In Heaven
- ⇒ Smokey Robinson & The Miracles, The Tracks Of My Tear
- ⇒ Van Halen, Jump
- ⇒ Pearl Jam, Jeremy
- ⇒ Bob Dylan, Tangled Up in Blue
- ⇒ Prince, Little Red Corvette

#### 11.1.6 Género rock.

- ⇒ Stairway to Heaven, Led Zeppelin
- ⇒ Hey Jude, The Beatles
- ⇒ You Can't Always Get What You Want, The Rolling Stones
- ⇒ Another Brick in the Wall, Pink Floyd
- ⇒ All Along the Watchtower, Jimi Hendrix
- ⇒ Hotel California, The Eagles
- ⇒ La Grange, ZZ Top
- ⇒ You Really Got Me, Van Halen
- ⇒ Layla, Derek And The Dominos
- ⇒ Riders on the Storm, The Doors
- ⇒ Won't Get Fooled Again, The Who
- ⇒ Bohemian Rhapsody, Queen
- ⇒ Dream On, Aerosmith
- ⇒ More Than a Feeling, Boston
- ⇒ Sultans of Swing, Dire Straits

- ⇒ Born on the Bayou, Creedence Clearwater Revival
- ⇒ Roundabout, Yes
- ⇒ Carry on Wayward Son, Kansas
- ⇒ Funeral for a Friend, Elton John
- ⇒ Locomotive Breath, Jethro Tull
- ⇒ I Still Haven't Found, U2
- ⇒ Magic Carpet Ride, Steppenwolf
- ⇒ Love Alive, Heart
- ⇒ Sweet Home Alabama, Lynyrd Skynyrd
- ⇒ Like a Rolling Stone, Bob Dylan
- ⇒ Purple Haze, Jimi Hendrix
- ⇒ Runnin' Down a Dream, Tom Petty
- ⇒ Let It Be, The Beatles
- ⇒ A Day in the Life, The Beatles
- ⇒ Every Breath You Take, The Police
- ⇒ Rock'n Me, Steve Miller
- ⇒ Brain Damage, Eclipse, Pink Floyd
- ⇒ Light My Fire, The Doors

- ⇒ Oh Well, Fleetwood Mac
- ⇒ Gold Dust Woman, Fleetwood Mac
- ⇒ Paranoid, Black Sabbath
- ⇒ Black Dog, Led Zeppelin
- ⇒ I've Seen All Good People, Yes
- ⇒ For What It's Worth, Buffalo Springfield
- ⇒ Black Magic Woman, Santana
- ⇒ Satisfaction, The Rolling Stones
- ⇒ You Shook Me, AC/DC
- ⇒ Bad Company, Bad Co.
- ⇒ In the Air Tonight, Phil Collins
- ⇒ Tush, ZZ Top
- ⇒ Time of the Season, The Zombies
- ⇒ Take the Long Way Home, Supertramp
- ⇒ Lonely is the Night, Billy Squire
- ⇒ My Best Friend's Girl, The Cars
- ⇒ Sweet Emotion, Aerosmith
- ⇒ Don't Fear the Reaper, Blue Oyster Cult

#### 11.1.7 Género salsa.

- ⇒ Periódico de ayer, Héctor Lavoe
- ⇒ Cali pachanguero, Niche
- ⇒ Gitana, Willie Colón
- ⇒ El cantante, Héctor Lavoe

- ⇒ Pedro Navaja, Rubén Blades
- ⇒ Llorarás, Dimensión Latina Oscar d'León
- ⇒ Timbalero, El gran Combo
- ⇒ Devórame otra vez, Lalo Rodríguez

- ⇒ Rebelión, Joe Arroyo
- ⇒ Sonido Bestial, Richie Ray & Bobby Cruz
- ⇒ Eldía de mi suerte, Héctor Lavoe
- ⇒ Plástico, Rubén Blades
- ⇒ Ven Morena, Oscar D'León
- ⇒ La fiesta de Pilito, El gran combo
- ⇒ LLuvia, Eddie Santiago
- ⇒ La Cura, Frankie Ruiz
- ⇒ Calle Luna, Calle Sol, Héctor Lavoe y Willie Colón
- ⇒ Desesperado, Luis Enrique
- ⇒ Mi libertad, Frankie Ruiz
- ⇒ Juanito Alimaña, Héctor Lavoe
- ⇒ Aquel viejo motel, David Pavón
- ⇒ El Amor, Ález León
- ⇒ Llegó, Croma Latina
- ⇒ Esperando Primavera, Croma Latina
- ⇒ Que bueno baila usted, Oscar D'León
- ⇒ Micaela, Sonora Carruseles
- ⇒ Brujería, El gran Combo
- ⇒ Otra oportunidad, Jimmy Bosch
- ⇒ Una Aventura, Grupo Niche
- ⇒ Desde que no éstas, Desde que no éstas
- ⇒ Al son de los cueros, Sonora Carruseles
- ⇒ Aicha, Africando
- ⇒ Algo eléctrico, Naty y su orquesta
- ⇒ Ave Maria Lola, Sonora Carruseles
- ⇒ Como les gusta, Son Candela
- ⇒ Ahora si, Sonora Ponceña
- ⇒ Cachondea, Fruko y sus Tesos
- ⇒ Azucar, Eddy McLean y Merensalsa
- ⇒ Ay no, Ray Barretto
- ⇒ De profesion... tu amante, Eddie Santiago
- ⇒ Así fue, Nelson Díaz y la constelación
- ⇒ Agua Que va a Caer, Ismael Rivera
- ⇒ Oye como va, Celia cruz
- ⇒ Lección de amor, adalberto álvarez
- ⇒ Bogna Sirala, Africando
- ⇒ Cada vez que te veo, Afro Cuban Allstars y felix Baloy
- ⇒ Amigo, Marc Anthony
- ⇒ Azucar, Celia Cruz
- ⇒ Procura, Chichi Peralta
- ⇒ Amores como el nuestro, Jerry Rivera

## 11.2 Pistas de Pruebas

### 11.2.1 Género clásica.

- ⇒ Sinfonía N° 40 (Wolfgang Amadeus Mozart) gran sinfonía en sol menor
- ⇒ Sinfonía N° 1 (Ludwig van Beethoven)
- ⇒ The Lark Ascending (Ralph Vaughan Williams)
- ⇒ Concierto Para Clarinete en La Mayor (Wolfgang Amadeus Mozart)
- ⇒ Fantasía Sobre un Tema de Thomas Tallis (Ralph Vaughan Williams)
- ⇒ Canon en re mayor - Johann Pachelbel
- ⇒ Concierto Para Piano y Orquesta N° 2 (Sergéi Rajmáninov)
- ⇒ Concierto Para Piano N° 2 (Frédéric Chopin)
- ⇒ Concierto de Aranjuez (Joaquín Rodrigo)
- ⇒ Variaciones Enigma (Edward Elgar)

### 11.2.2 Género Electrónica.

- ⇒ White Noise (feat. AlunaGeorge), Disclosure
- ⇒ Bubble Butt (Radio Mix) (feat. Bruno Mars, 2 Chainz, Tyga & Mystic), Major Lazer
- ⇒ Bad Girls, M.I.A.
- ⇒ Crave You (Adventure Club Remix), Flight Facilities
- ⇒ Promises (Skrillex & Nero Remix), Nero
- ⇒ Crystallize, Lindsey Stirling
- ⇒ Paper Planes, M.I.A.
- ⇒ One More Time, Daft Punk
- ⇒ Latch (feat. Sam Smith), Disclosure
- ⇒ Chainsaw (feat. Tedashii), Family Force 5

### 11.2.3 Género Hip-Hop.

- ⇒ Same Love (feat. Mary Lambert), Macklemore & Ryan Lewis
- ⇒ Can't Hold Us (feat. Ray Dalton), Macklemore & Ryan Lewis
- ⇒ Wop (Official Version), J. Dash
- ⇒ Thrift Shop (feat. Wanz), Macklemore & Ryan Lewis
- ⇒ Holy Grail (feat. Justin Timberlake), JAY Z

- ⇒ Can't Hold Us (feat. Ray Dalton), Macklemore & Ryan Lewis
- ⇒ Are You Getting Enough? (feat. Miles Kane), Professor Green
- ⇒ Ring a Ling, Sneakbo
- ⇒ Thrift Shop (feat. Wanz), Macklemore & Ryan Lewis

- ⇒ Lights On (feat. Angel & Tinchy Stryder), Wiley
- ⇒ Wild Ones (feat. Sia), Flo Rida
- ⇒ Hey Porsche, Nelly
- ⇒ My Time (feat. Bridget Kelly), K Koke
- ⇒ Heatwave (feat. Ms D), Wiley
- ⇒ Mama Do the Hump, Rizzle Kicks

#### 11.2.4 Género Jazz.

- ⇒ Feeling Good, Nina Simone
- ⇒ L-O-V-E, Nat "King" Cole
- ⇒ Georgia On My Mind, Ray Charles & The Count Basie Orchestra
- ⇒ La Vie En Rose (Single Version), Louis Armstrong
- ⇒ Lucky, Kat Edmonson

- ⇒ Green Onions, Booker T. & The M.G.'s
- ⇒ In the Mood, Glenn Miller
- ⇒ Songbird, Kenny G
- ⇒ Wake Me Up (A Tribute to Avicii and Aloe Blacc), Spark Productions
- ⇒ (You Don't Know) How Glad I Am, Nancy Wilson

#### 11.2.5 Género Pop.

- ⇒ Best Song Ever, One Direction
- ⇒ We Can't Stop, Miley Cyrus
- ⇒ Baby I, Ariana Grande
- ⇒ Get Lucky (feat. Pharrell Williams), Daft Punk
- ⇒ Burn, Ellie Goulding

- ⇒ Slow Down, Selena Gomez
- ⇒ Love Somebody, Maroon 5
- ⇒ Cruise (Remix) (feat. Nelly), Florida Georgia Line
- ⇒ The Other Side, Jason Derulo
- ⇒ Mirrors, Justin Timberlake

### 11.2.6 Género Rock.

- ⇒ Wildfire, John Mayer
- ⇒ Hail to the King, Avenged Sevenfold
- ⇒ All Summer Long, Kid Rock
- ⇒ Don't Stop Believin, Journey
- ⇒ Picture (feat. Sheryl Crow), Kid Rock
- ⇒ I'll Follow You, Shinedown
- ⇒ I Got You, Jack Johnson
- ⇒ Start Me Up (Hyde Park Live / 2013), The Rolling Stones
- ⇒ Drops of Jupiter, Train
- ⇒ Because We Can, Bon Jovi

### 11.2.7 Género Salsa.

- ⇒ Marc Anthony, Vivir Mi Vida
- ⇒ Alex Matos, Amor Perfecto
- ⇒ Chiquito Team band, Volvere
- ⇒ Felix Manuel, Dos Amantes
- ⇒ Proyecto A, Renuncio
- ⇒ Michel, No me compares
- ⇒ Yiyo Sarante, En Bandolera
- ⇒ Sexappeal, Esta Noche
- ⇒ Antonio Cartagena, Sin Ti
- ⇒ Victor Wailly, Si te hubiera conocido ayer

### 11.3 Referencias de entrenamiento masivo

Para algunas pruebas se decidió incrementar el número de muestras de entrenamiento de 50 a 500. Por practicidad no se enlistan las canciones, sin embargo a continuación se muestran las referencias de dónde se obtuvieron las pistas de cada género.

- ⇒ Clásica: ClassicalArchives, The ultimate classical music destination. (<http://www.classical-archives.com/index.html>)
- ⇒ Electrónica: Bleep, ELECTRONIC & ELECTRONICA. <https://bleep.com/stream-/ELECTRONIC%20&%20ELECTRONICA>
- ⇒ Hip-hop:Rate Your Music. The 500 Greatest Hip Hop albums plus the other ones that are honorable mention. [http://rateyourmusic.com/list/ChrisPCthe\\_500\\_greatest\\_hip\\_hop\\_albums\\_plus\\_the\\_other\\_ones\\_that\\_are\\_honorable\\_mention](http://rateyourmusic.com/list/ChrisPCthe_500_greatest_hip_hop_albums_plus_the_other_ones_that_are_honorable_mention)
- ⇒ Jazz: Sony Music, The Perfect Jazz Collection - 25 Original Albums. <http://www.-sonymusic.de/VariouS/La-Discotheque-Ideale-En-25-Albums-Originiaux/P/2383936>
- ⇒ Pop: MTV, Rolling Stone list top 100 pop songs since 1963. [http://www2.ljworld.com-/news/2000/nov/19/mtv\\_rolling\\_stone/](http://www2.ljworld.com-/news/2000/nov/19/mtv_rolling_stone/)
- ⇒ Rock: Q104.3, The Top 1,043 Classic Rock Songs of All Time: Dirty Dozenth Edition [http://www.q1043.com/common/top\\_songs/2012.html](http://www.q1043.com/common/top_songs/2012.html)
- ⇒ Salsa: POSTGORDI, Top 500 Salsa Of All Time + Top 500 Salsa Songs. <http://-postgordi.blogspot.mx/2012/11/descarga-las-mejores-500-canciones-de.html>

# Referencias

- [1] Iñaki Inza y Pedro Larrañaga Abdelmalik Moujahid. Clasificadores k-nn, Departamento de ciencias de la Computación e Inteligencia Artificial, Universidad del País Vasco-Euskal Herriko Unibertsitatea.
- [2] Joachim E. Berendt. El jazz. de nueva orleans al jazz rock. fondo de cultura económica., 2003.
- [3] Kun-Ming Yu Chang-Hsing Lee, Jau-Ling Shih. Automatic music genre classification using modulation spectral contrast feature, 2007.
- [4] dedicadisimo2003.webcindario.com. Música electrónica., 2007.
- [5] Min Xu et al. Generación de palabras clave de audio basado en hmm, 2004.
- [6] La evolución de la música pop. Centros 4. La evolución de la música pop, 2010.
- [7] generosmusicales.supaw.com Géneros musicales. Géneros musicales, 2013.
- [8] Sergio Suárez Guerra. Cálculo lpc y cepstrum, 2013.
- [9] Emilia Gómez Gutierrez. Digitalización del sonido, 2009.
- [10] [www.juntadeandalucia.es/averroes/iesmateoaleman /musica / elgeneromusical.htm](http://www.juntadeandalucia.es/averroes/iesmateoaleman/musica/elgeneromusical.htm) Los géneros musicales. Los géneros musicales, abril 2013.
- [11] International Federation of the Phonographic Industry. Digital music report 2013, 2013.
- [12] Goutam Sahidullah, Saha. Diseño, análisis y evaluación experimental de la transformación basada en bloques en el cálculo mfcc para el reconocimiento del altavoz, 2012.
- [13] [scikit learn.org](http://scikit-learn.org). Machine learning in python, scikit-learn 0.14 stable.

- [14] J. Shepherd. Continuum encyclopedia of popular music of the world: Performance and production, 2003.
- [15] [en.wikipedia.org/wiki/ Cepstrum](http://en.wikipedia.org/wiki/Cepstrum) Wikipedia. Cepstrum., 2013.
- [16] [en.wikipedia.org/wiki/ Digital\\_audio](http://en.wikipedia.org/wiki/Digital_audio) Wikipedia. Digital audio, 2013.
- [17] [en.wikipedia.org/wiki/Mel-frequency\\_cepstrum](http://en.wikipedia.org/wiki/Mel-frequency_cepstrum) Wikipedia. Mel-frequency cepstrum., 2013.
- [18] [en.wikipedia.org/wiki/ Rock\\_music](http://en.wikipedia.org/wiki/Rock_music) Wikipedia. Rock music, 2013.
- [19] [es.wikipedia.org/wiki / Sonido](http://es.wikipedia.org/wiki/Sonido) Wikipedia. Sonido, 2013.
- [20] [es.wikipedia.org/ Género\\_musical](http://es.wikipedia.org/wiki/G%C3%A9nero_musical) Wikipedia. Sonido, 2013.
- [21] [es.wikipedia.org/wiki/ Hip\\_hop](http://es.wikipedia.org/wiki/Hip_hop) Wikipedia. Hip hop., 2013.
- [22] [es.wikipedia.org/wiki/ Música\\_electrónica](http://es.wikipedia.org/wiki/M%C3%BAsica_electr%C3%B3nica) Wikipedia. Música electrónica., 2013.
- [23] Xin Zhang and Zbigniew W. Ras. Analysis of sound features for music timbre recognition, 2007.

## Conclusiones

Al presentar la propuesta de este trabajo se investigó sobre características que permitieran o lograran clasificar el sonido de alguna forma, se encontraron trabajos y proyectos en los que se obtienen muy buenos resultados reconociendo la voz de individuos con características en el dominio de la frecuencia aunque también en algunos otros se utilizaban características en el dominio del tiempo. También existen trabajos en los que se logra diferenciar el sonido de distintos instrumentos musicales e incluso la nota que se está reproduciendo, para estos trabajos se utilizan características en el dominio de la frecuencia.

Para iniciar el trabajo de investigación se comenzó a trabajar con características en el dominio del tiempo, que según investigación realizada lograban identificar voz humana. También se propuso utilizar dos clasificadores, una máquina de soporte vectorial y un clasificador de vecinos más cercanos. Se propuso un número de 50 muestras de entrenamiento por cada género (un total de 350 muestras de entrenamiento) y 10 muestras para pruebas por cada género (70 muestras de prueba). Y así fue como se realizó la primer prueba integral de clasificación en la que se obtuvieron muy malos resultados pues ambos clasificadores al momento de hacer su predicción siempre devolvían el mismo género sin importar el género de la muestra de prueba, al ver estos resultados se comenzó a afinar el proceso de extracción de características con el fin de estandarizar las muestras y que no afectaran condiciones como el volumen con el que son grabadas las pistas musicales. Así se volvieron a realizar pruebas en las que se observó un mejor desempeño de los clasificadores pero aún sin ser óptimo.

Al no obtener los resultados esperados con las características en el dominio del tiempo comenzamos a utilizar las características en el dominio de la frecuencia. Con estas características se mejoraron los resultados de los clasificadores, se realizaron pruebas haciendo distintas combinaciones entre el número de características que se obtienen y la forma en la que se obtienen y también en los parámetros que reciben los clasificadores. Se pudo observar que las frecuencias de los géneros presentan demasiada similitud, sin embargo es

posible identificar al género del rock con el de hip-hop como el par que muestra menos similitud. Por este motivo se decide incrementar el número de muestras de entrenamiento de estos dos géneros de 50 a 500 muestras. De esta manera se mejoraron los resultados de clasificación entre estos dos géneros con características en el dominio de la frecuencia.

Se continúa haciendo pruebas de eficiencia de los clasificadores para encontrar sus mejores parámetros y los vectores con las características más óptimas para incrementar el número de predicciones acertadas de nuestro clasificador. En conjunto se está desarrollando un proceso de clasificación por medio de un árbol de decisiones que implemente los mejores parámetros y vectores de características para reconocer cada par de géneros.

Resulta complejo encontrar características que logren una óptima clasificación de pistas musicales de acuerdo a su género, los resultados obtenidos hasta ahora son producto de las pruebas realizadas haciendo diferentes combinaciones entre la manera en la que operan los algoritmos de clasificación del KNN y SVM, de los vectores de características y de la forma en la que se extraen las componentes de la señal de audio que generan las características, obteniendo así un gran número de combinaciones de pruebas para lograr una clasificación. Esta tarea se complica más si tomamos en cuenta que al analizar una señal de audio de una pista musical es posible encontrar una o más voces, el volumen de cada voz, el uso de uno o más instrumentos, el volumen con el que suena cada instrumento, el momento y tiempos en los que suena cada voz o instrumento, el patrón con el que se escucha cierto instrumento a cierto volumen y en alguna nota musical, y así podríamos seguir numerando variables que pudieran ser analizadas de muchas formas y pudiendo no encontrar un patrón que distinga el género musical.