



CATEGORIZACIÓN Y RECONOCIMIENTO DE EMOCIONES EN PIEZAS MUSICALES MEDIANTE APRENDIZAJE AUTOMÁTICO.

TRABAJO FIN DE GRADO
GRADUADO EN INGENIERÍA INFORMÁTICA
ETSIINF UPM

Autor: Rubén Ibáñez Redondo



OBJETIVOS

- Analizar piezas musicales usando algoritmos de Aprendizaje Automático.
- Buscar la forma de agrupar (clusterizar) un conjunto de piezas musicales en categorías.
- Implementar un mecanismo de aprendizaje supervisado (clasificación) para reconocer las emociones transmitidas por una pieza.
- Buscar una representación de los datos que pueda ser utilizada como input de estos algoritmos.

ÍNDICE

Desarrollo

Investigación
formatos
representación

Investigación
Datasets
existentes

Elección Software

Elección y
preparación
Dataset

Análisis Inicial WEKA

Preprocesamiento

Wrappers

Conclusiones

Clasificación Supervisada

IBk, J48, JRip,
NaiveBayes y
TAN.

Análisis
resultados
algoritmos

Comparación
resultados con
Wrapper

Clasificación No Supervisada

Técnica
Jerárquica

Técnica
Particional

Técnica
Probabilística

Relación Emociones - Géneros

Recopilar datos
manualmente

Añadir
'emociones'
como atributo

Añadir
'emociones'
como atributo
clase

Comparación y
conclusiones

DESARROLLO

- Representación de una pieza musical (objeto complejo) para hacer clustering.
- Extracción de atributos más comunes (los acordes, las armonías, la melodía, el tono principal, los compases por minuto o el ritmo en la pieza).
- Formato MIDI y Million Song Dataset.
- WEKA - Plataforma de software para el aprendizaje automático y la minería de datos escrita en Java.

DATASET: MUSIC FEATURES

- Contenido: características extraídas a partir de un conjunto de datos que consta de 1000 pistas de audio cada 30 segundos de duración.
- Contiene 10 géneros, cada uno representado por 100 pistas.
- Las pistas son archivos de audio Mono de 16 bits de 22050Hz en formato .wav

- 30 atributos

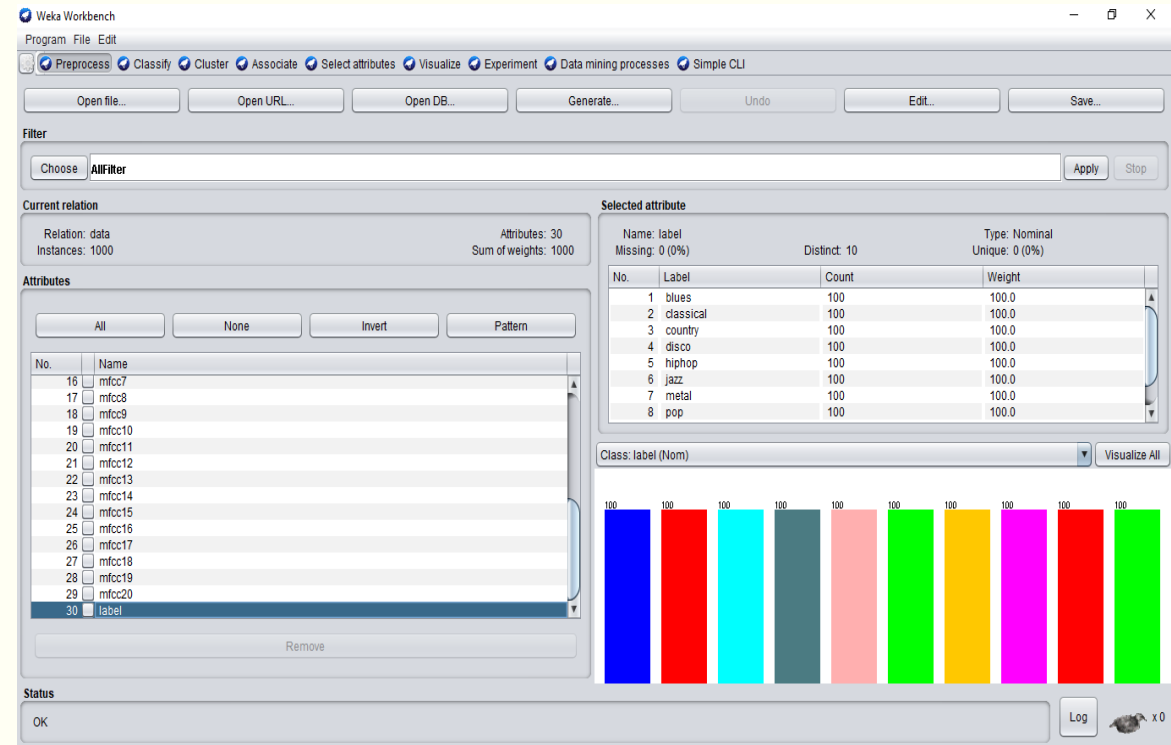
label	
tempo	rolloff
beats	zero_crossing_rate
chroma_stft	20 coeficientes cepstrales
rmseRoot	spectral_bandwidth
spectral_centroid	filename

- Atributo clase representa el género

Blues	Jazz
Classical	Metal
Country	Pop
Disco	Reggae
Hiphop	Rock

ANÁLISIS WEKA

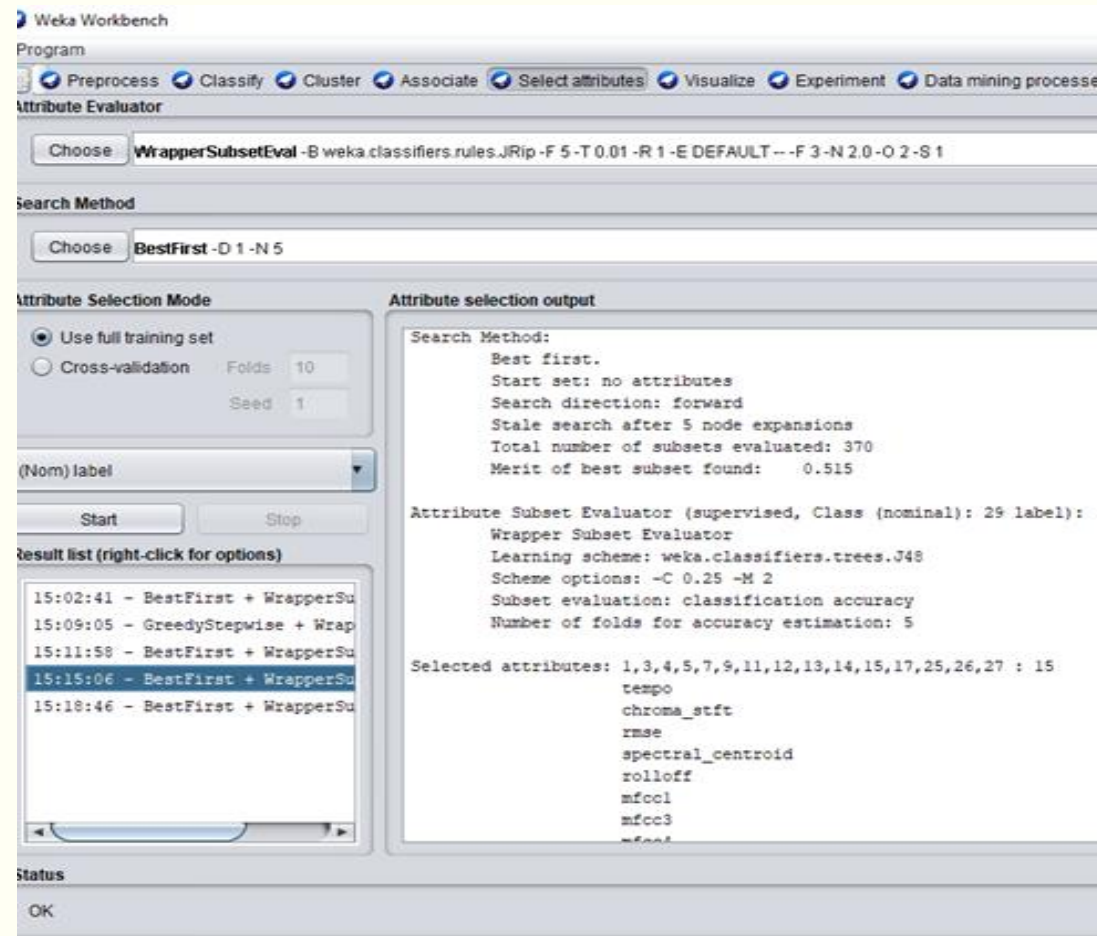
- Preprocesamiento: filtros para la discretización, normalización, reemplazamiento y combinación de atributos.
- Reducir el número de atributos.
- Mejorar el rendimiento de un algoritmo de aprendizaje.
- Distintos métodos.
- Conclusión. Atributos más relevantes: tempo, beats, chroma_stft, rmae, special_centroid, spectral_bandwidth, rolloff y zero_crossing_rate.



Dataset cargado en WEKA

WRAPPER

- Evaluador de subconjuntos de atributos mediante validación cruzada.
- Usará siempre un clasificador.
- En función del modelo de clasificación, resultados de subconjuntos de entre 7 y 23 atributos.
- Variedad en cada uno y difícil encontrar un patrón.
- Solución: Hacer una tabla con los resultados específicos del Wrappers para cada clasificador.



Resultados Wrapper con algoritmo JRip y BestFirst

CLASIFICACIÓN SUPERVISADA

- Encontrar propiedades comunes entre un conjunto de datos y clasificarlos dentro de diferentes clases, de acuerdo con un modelo de clasificación.
- Algoritmos de aprendizaje:
 - 1) J48: método expresados a través de árboles de decisión.
 - 2) IB1: basado en criterios de vecindad.
 - 3) IBk: $k=5$ número de vecinos más cercanos.
 - 4) Naive Bayes: clasificador probabilístico fundamentado en el teorema de Bayes.
 - 5) TAN: clasificador bayesiano simple aumentado con un árbol.
 - 6) JRip: construye un conjunto de reglas y las optimiza.

CLASIFICACIÓN SUPERVISADA

- Tabla comparación resultados clasificadores:

Clasificador	Modo de prueba	Instancias bien clasificadas (%)	Instancias mal clasificadas (%)	Error absoluto medio	Error absoluto relativo (%)
IB1	Validación cruzada	60.5	39.5	0.0801	44.5036
IBk (k=5)	Validación cruzada	62.4	37.6	0.0958	53.2295
J48	Validación cruzada	47.5	52.5	0.1083	60.1766
JRip	Validación cruzada	47.4	52.6	0.1216	67.5295
Naive Bayes	Validación cruzada	43.8	56.2	0.1146	63.6762
TAN	Validación cruzada	43.5	56.5	0.114	63.3298

CLASIFICACIÓN SUPERVISADA - WRAPPER

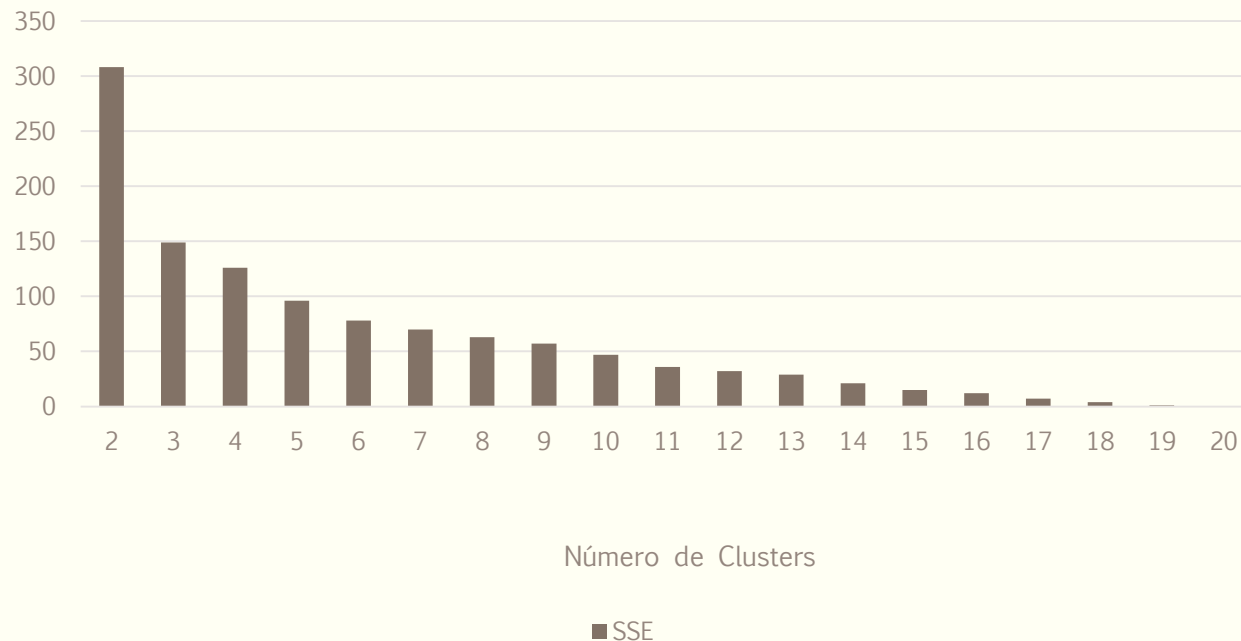
Clasificador	Instancias bien clasificadas (%)	Instancias mal clasificadas (%)	Error absoluto medio	Error absoluto relativo (%)
IB1	60.5	39.5	0.0801	44.5036
IB1 Wrapper BestFirst	63.2	36.8	0.0748	41.537
IB1 Wrapper GreedyStepwise	62.3	37.7	0.0767	42.6365
IBk (k=5)	62.4	37.6	0.0958	53.2295
IBk (k=5) Wrapper BestFirst	66	34	0.0923	51.2599
IBk (k=5) Wrapper GreedyStepwise	60.8	39.2	0.1003	55.7056
J48	47.5	52.5	0.1083	60.1766
J48 Wrapper BestFirst	50.7	49.3	0.1036	57.5579
J48 Wrapper GreedyStepwise	46.7	53.3	0.1137	63.1428
Naïve Bayes	43.8	56.2	0.1146	63.6762
Naïve Bayes Wrapper BestFirst	51.3	48.7	0.1085	60.2669
Naïve Bayes Wrapper GreedyStepwise	51.3	48.7	0.1085	60.2669

CLASIFICACIÓN NO SUPERVISADA

- Encontrar grupos naturales en un conjunto de datos muy grande no etiquetados y sin una clase.
- Homogeneidad dentro de las clases y heterogeneidad entre las distintas clases.
- Tres técnicas existentes y su algoritmo representativo:
 - 1) Jerárquica: SimpleKMeans
 - 2) Particional: HierarchicalClusterer.
 - 3) Probabilística: Expectation-maximization (EM).

SIMPLEKMEANS

- Elección de 10 clusters
- 26 iteraciones

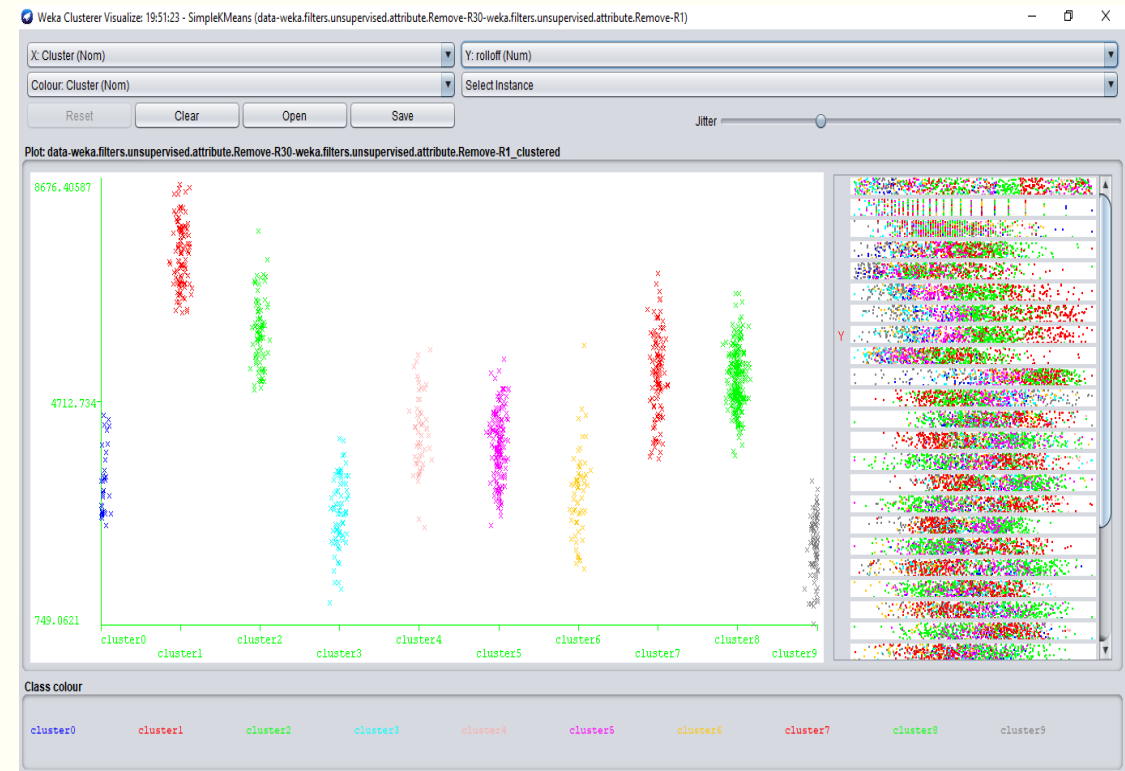


■ N° instancias:

1. 33 (3%)
2. 109 (11%)
3. 82 (8%)
4. 75 (8%)
5. 55 (6%)
6. 156 (16%)
7. 61 (6%)
8. 107 (11%)
9. 232 (23%)
10. 90 (9%)

SIMPLEKMEANS

- Análisis centroides:
 - Cluster 0: 'tempo' y 'beats' muy altos.
 - Cluster 1: 'spectral_bandwidth', y 'spectral_centroid' 'rolloff', más altos.
 - Cluster 3: 'tempo' y 'beats' muy bajos.
 - Cluster 5: muy neutro.
 - Cluster 6: valores de media muy bajos.
 - Cluster 8: 'chroma_stft' muy alto.



Cluster 1 – Valores atributo 'rolloff'

SIMPLEKMEANS

■ Análisis instancias Clusters:

- ❖ Cluster 0: Classical
- ❖ Cluster 1: Pop
- ❖ Cluster 2: Hiphop
- ❖ Cluster 3: Blues
- ❖ Cluster 4: Reggae
- ❖ Cluster 5: Country/Rock
- ❖ Cluster 6: Blues/Country
- ❖ Cluster 7: *Indefinido*
- ❖ Cluster 8: Metal
- ❖ Cluster 9: Classical

□ Cluster 1:

- Country: 1 instancia
- Disco: 20 instancias
- Hiphop: 9 instancias
- Jazz: 4 instancias
- Metal: 1 instancia
- **Pop: 64 instancias**
- Reggae: 7 instancias
- Rock: 3 instancias

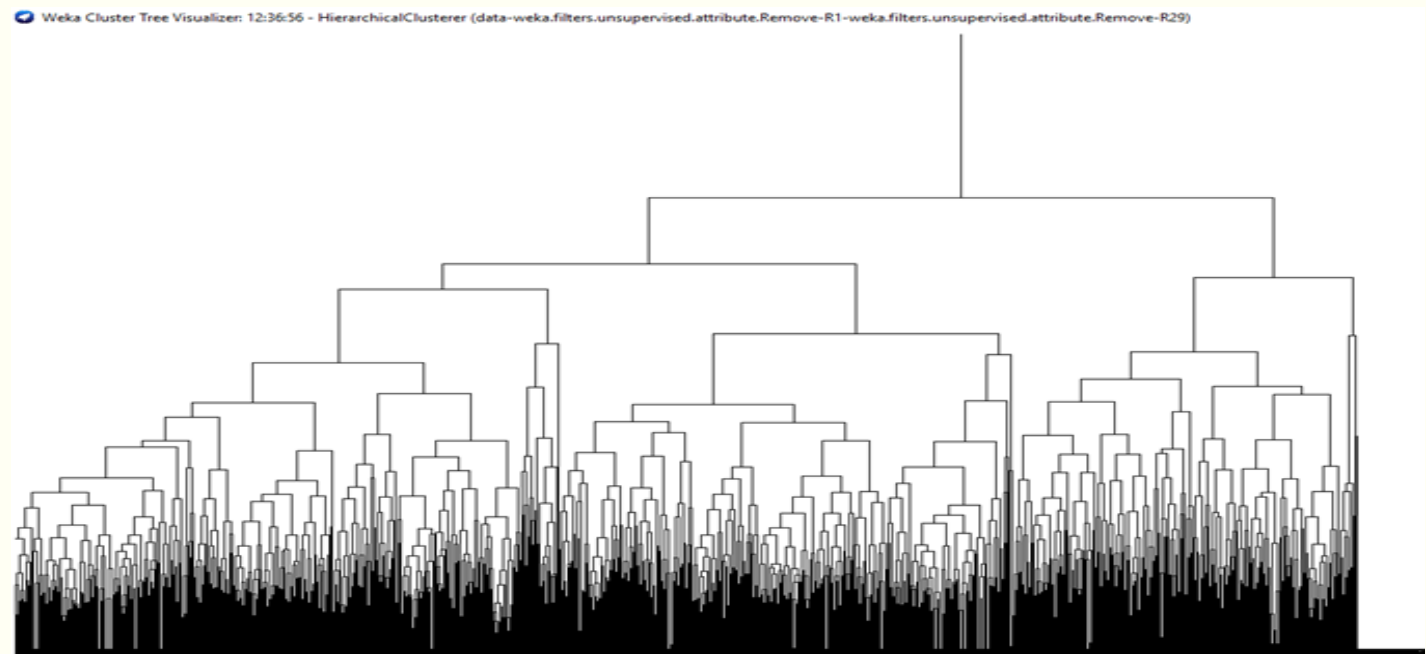
□ Cluster 7:

- Blues: 3 instancias
- Country: 20 instancias
- Disco: 12 instancias
- Hiphop: 14 instancias
- Jazz: 14 instancias
- Pop: 15 instancias
- Reggae: 10 instancias
- Rock: 19 instancias

HIERARCHICAL CLUSTERER

- Su salida no proporciona mucha información.
- Análisis de Dendrogramas.
- Distintos métodos de enlazar clusters:

- SINGLE
- COMPLETE
- AVERAGE
- CENTROID
- WARD



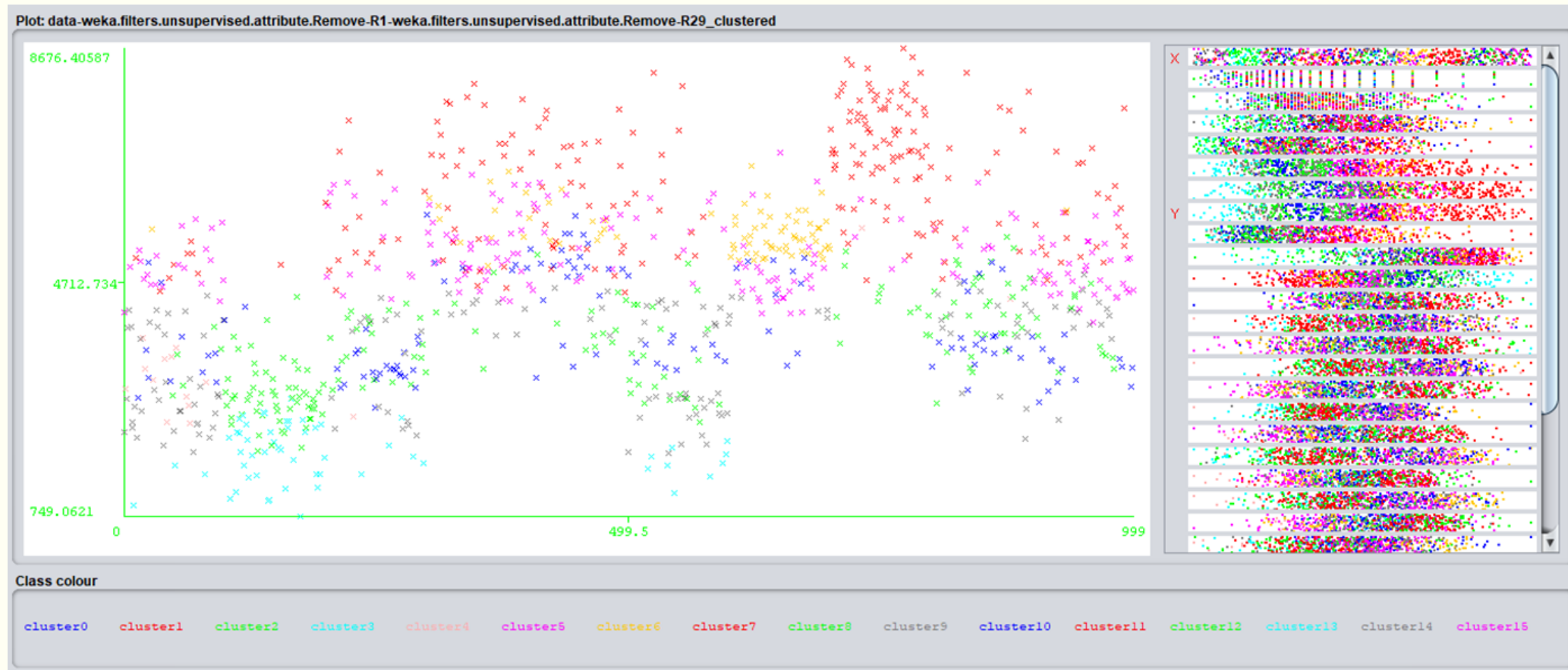
Dendrograma enlace completo o vecino más lejano (COMPLETE)

EXPECTATION – MAXIMIZATION (EM)

- Dada una muestra de datos, la probabilidad de cada instancia de pertenecer a un determinado cluster.
- 16 clusters mediante validación cruzada.
- 97 iteraciones.
- Clasificación de instancias:

▪ 0	44 (4%)	▪ 5	65 (7%)	▪ 10	79 (8%)
▪ 1	122 (12%)	▪ 6	77 (8%)	▪ 11	87 (9%)
▪ 2	64 (6%)	▪ 7	9 (1%)	▪ 12	63 (6%)
▪ 3	38 (4%)	▪ 8	53 (5%)	▪ 13	18 (2%)
▪ 4	15 (2%)	▪ 9	80 (8%)	▪ 14	61 (6%)
				▪ 15	125 (13%)

EXPECTATION – MAXIMIZATION (EM)



- Gráfico valores de las medias de los atributos para cada posible cluster.

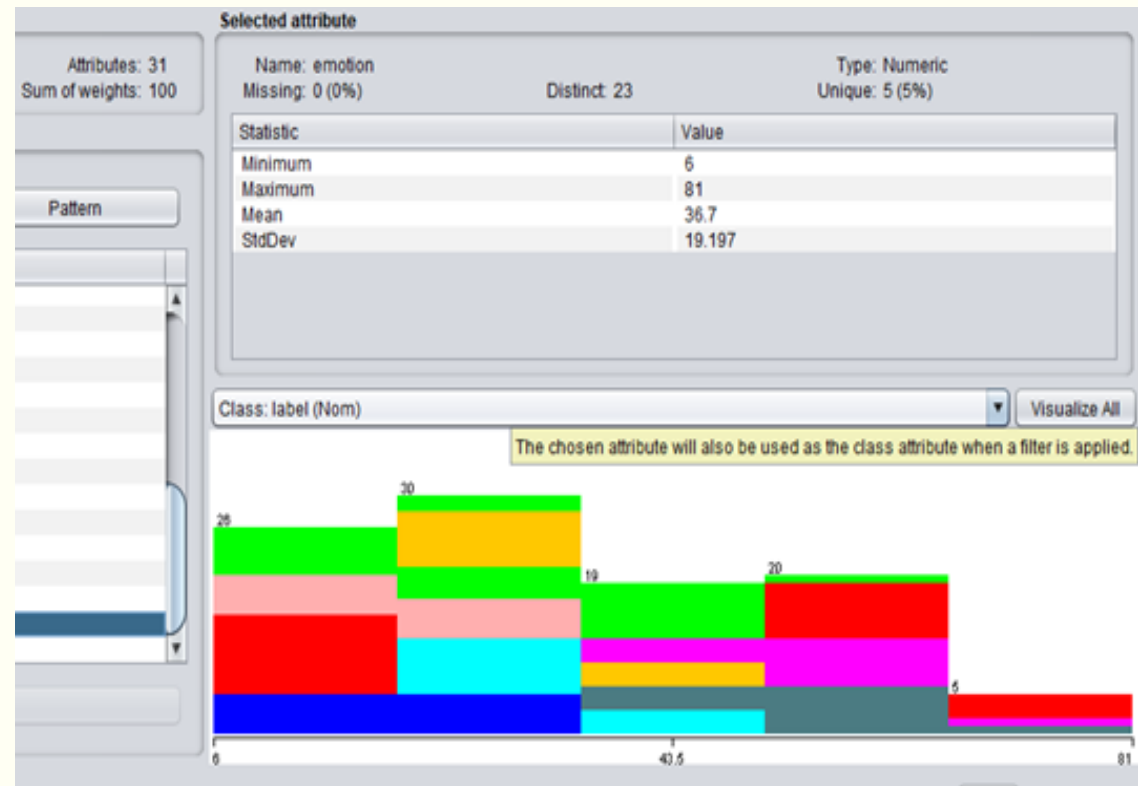
RELACIÓN EMOCIONES – GÉNEROS

- Medir emociones que le provocan a la gente canciones de nuestro dataset.
- Creación nuevo atributo para cuantificar esas emociones.
- Tomamos 10% como muestra y testeamos sobre distintas personas.
- ‘Feeling’ del 1 al 10 y ‘energy’ del 1 al 10. Multiplicamos resultados y creación nuevo atributo ‘emotion’. Valores del 1 al 100.
- Inclusión del atributo ‘emotion’ en nuestro dataset como un atributo más.

RELACIÓN EMOCIONES – GÉNEROS

- Valor medio 'emotion' en cada género:

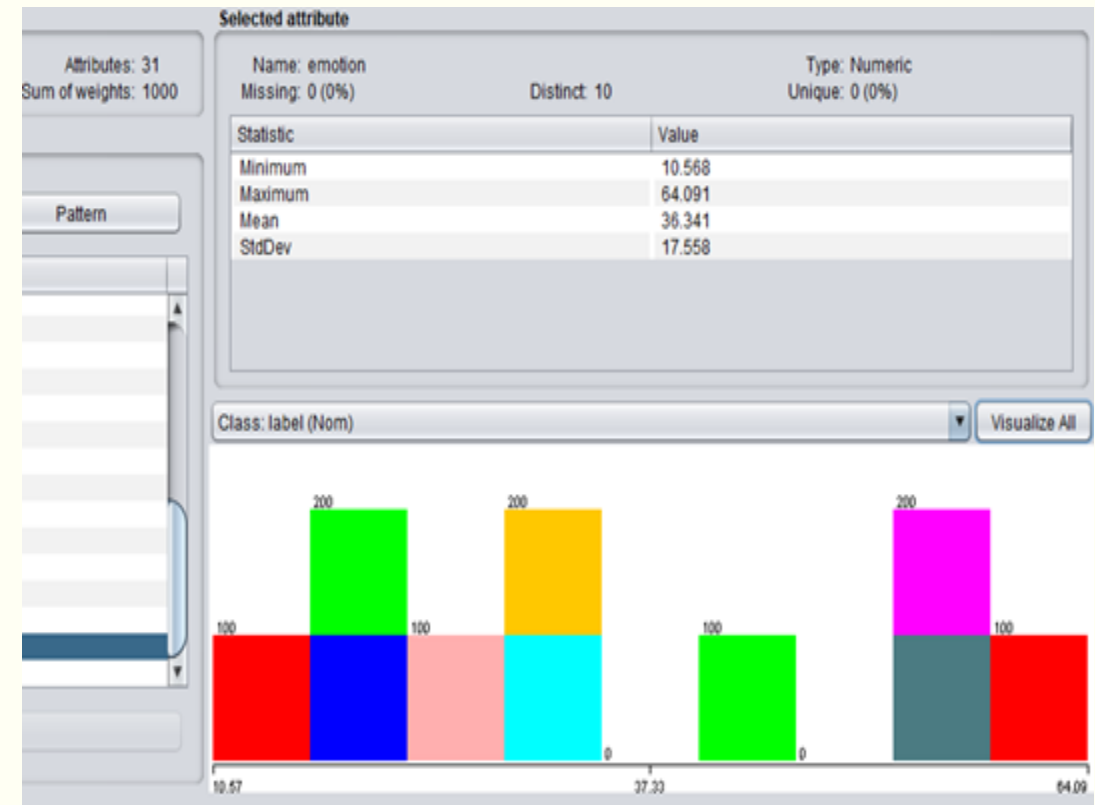
- Blues: 20.1136
- Classical: 10.5682
- Country: 33.0682
- Disco: 57.6136
- Hiphop: 23.1818
- Jazz: 20.4545
- Metal: 33.75
- Pop: 57.2727
- Reggae: 64.0909
- Rock: 43.2955



Información atributo 'emotion' Dataset 100 instancias.

RELACIÓN EMOCIONES – GÉNEROS

- Aplicamos lbk con KNN = 5 al Dataset de muestra y 49% instancias bien clasificadas.
- Extrapolamos los resultados a las 900 instancias restantes.
- Método: asignando a cada instancia el valor medio del cluster al que pertenece.
- Aplicamos lbk con KNN = 5 y 54.4% instancias bien clasificadas.



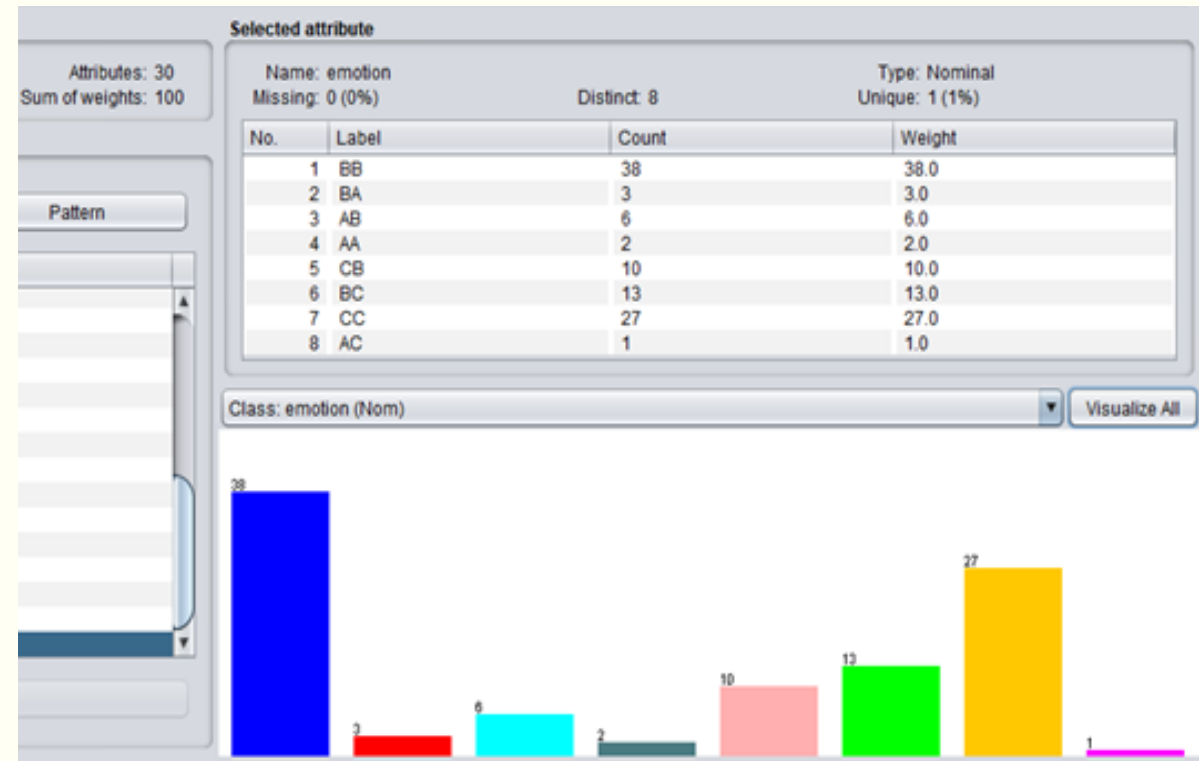
Información atributo 'emotion' Dataset 1000 instancias

RELACIÓN EMOCIONES – GÉNEROS

- Incluir atributo 'emotion' como atributo clase del Dataset.
- Comparamos valores por separado de 'feeling' y 'energy':
 - **Feeling:**
 - ❖ Minimun: 2
 - ❖ Maximun: 9
 - ❖ Mean: 5.78
 - ❖ StdDev: 1.661
 - **Energy:**
 - ❖ Minimun: 2
 - ❖ Maximun: 9
 - ❖ Mean: 6
 - ❖ StdDev: 1.735
- Nuevo método creación 'emotion': División valores atributos en 3 grupos (A:1-3, B:4-6 y C:7-9).
- Se juntan las 2 letras y 9 posibles nuevos clusters:
AA,AB,AC,BA,BB,BC,CA,CB,CC

RELACIÓN EMOCIONES – GÉNEROS

- Distribución instancias en los nuevos clusters:
 - AA – 2%
 - AB – 6%
 - AC – 1%
 - BA – 3%
 - BB – 38%
 - BC – 13%
 - CB – 10%
 - CC – 27%
- Aplicamos lbk con KNN = 5 al Dataset de muestra y 35% instancias bien clasificadas.



Información nuevo atributo clase 'emotion'.



FIN

GRACIAS POR SU ATENCIÓN