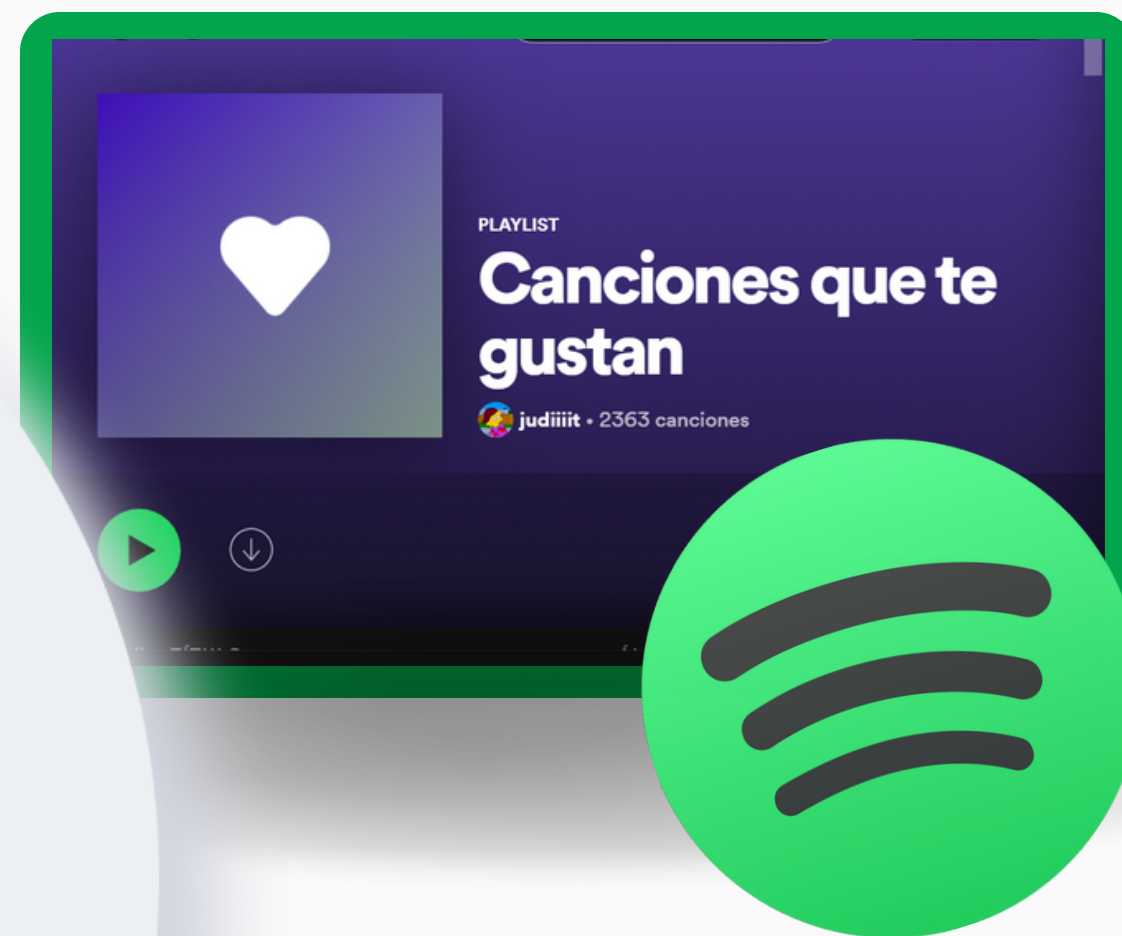
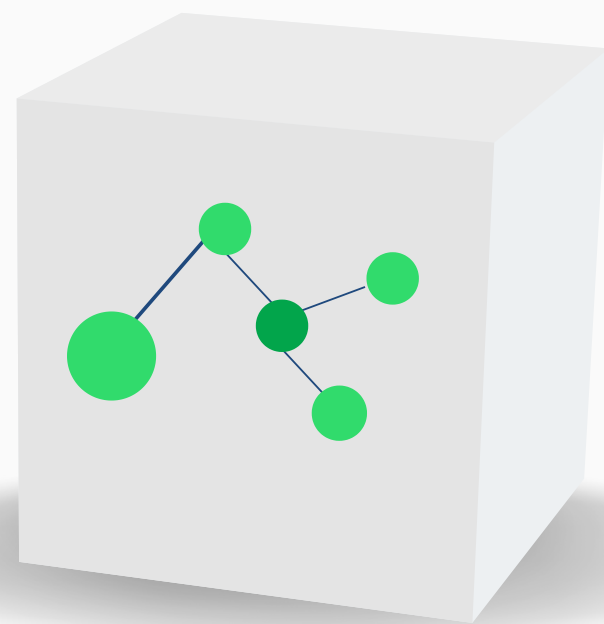


# *Estudio comparativo de precisión y explicabilidad en algoritmos de cajas blancas, negras y grises sobre modelos de recomendación musical*

---



*Judit González Prol*



Preparación  
del conjunto  
de datos

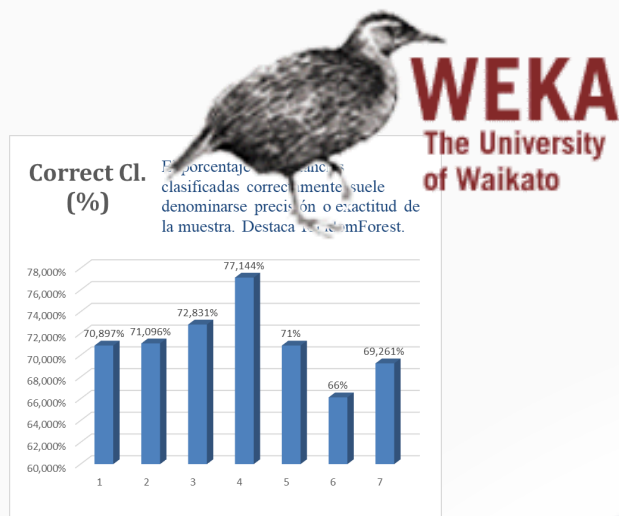
Obtención de  
las métricas  
de precisión  
con Weka

Desarrollo de  
explicaciones  
en lenguaje  
natural con  
Expliclas

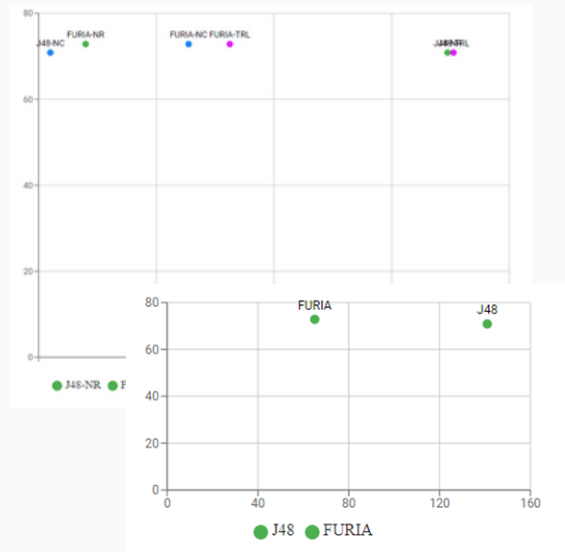
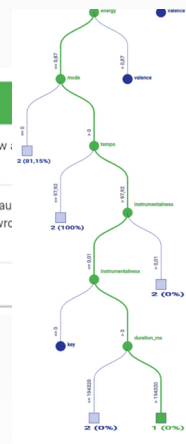
Análisis de  
precisión vs.  
explicabilidad

Determinación  
del algoritmo/s  
con mejores  
resultados

	A	B	C	D	E	F	G	H	I	J	K	L
1	id	acousticness	danceability	duration_ms	energy	instrumentalness	key	liveness	loudness	mode	speechiness	tempo
2	0.0	0.0102	0.433	204600.0	0.434	0.0219	2.0	165	-8.795	1.0	0.431	150.962
3	1.0	1.99	0.743	326933.0	0.359	0.00611	1.0	137	-10.401	1.0	0.0794	160.083
4	2.0	0.0344	0.838	185707.0	0.412	0.000334	2.0	159	-7.148	1.0	0.289	75.044
5	3.0	0.004	0.494	199413.0	0.338	0.51	5.0	0.022	-15.236	1.0	0.0261	96.498
6	4.0	0.18	0.478	392993.0	0.563	0.112	5.0	0.493	-11.648	0.0	0.0094	174.004
7	5.0	0.00479	0.804	251333.0	0.56	0.0	0.0	0.164	-6.682	1.0	0.185	85.023
8	6.0	0.0145	0.739	241400.0	0.472	7.27e-06	1.0	207	-11.204	1.0	0.156	80.03
9	7.0	0.0202	0.266	349687.0	0.348	0.64	10.0	16	-11.409	0.0	0.0371	144.154
10	8.0	0.0481	0.403	202853.0	0.944	0.0	0.1	0.342	-3.606	0.0	0.347	139.035
11	9.0	0.00208	0.836	226840.0	0.603	0.0	0.7	0.571	-7.792	1.0	0.237	99.994
12	10.0	0.0572	0.525	358187.0	0.855	0.0143	5.0	0.649	-7.372	0.0	0.0548	111.951
13	11.0	0.0915	0.753	324880.0	0.748	0.00348	10.0	212	-8.821	0.0	0.0494	104.122
14	12.0	0.251	0.403	356977.0	0.434	0.0619	0.0	108	-11.093	1.0	0.0342	127.681
15	13.0	0.366	0.762	243270.0	0.476	0.0	0.0	0.103	-12.686	1.0	0.114	130.007
16	14.0	0.44	0.662	247288.0	0.603	0.0	0.9	0.0972	-8.317	0.0	0.0793	125.011
17	15.0	0.019	0.437	188133.0	0.832	0.0968	0.0	0.316	-6.637	1.0	0.183	99.984
18	16.0	0.0239	0.601	270827.0	0.955	0.0451	1.0	119	-4.111	1.0	0.0458	123.922
19	17.0	0.233	0.789	447907.0	0.659	0.00049	4.0	184	-12.654	0.0	0.0429	122.415



Algorithm	Local explanation
J48	Example is 1 because acousticness, duration_ms, instrumentalness, loudness and mode are low : energy and tempo are medium.
FURIA	We have a high confidence in the classification result. It is very likely that this example is 2 becau loudness and liveness are low and instrumentalness and energy are medium. However, this is wr because the type should be 1 instead of 2 according to the information in the dataset.
RandomForest	This instance is classified as 1



# Fases del proyecto

# Algoritmos

Mayor  
interpretabilidad

Menor precisión

**Cajas  
Blancas**

- J48
- JRip

Cajas  
Grisas

- Bayes  
NET  
FURIA

**Cajas  
Negras**

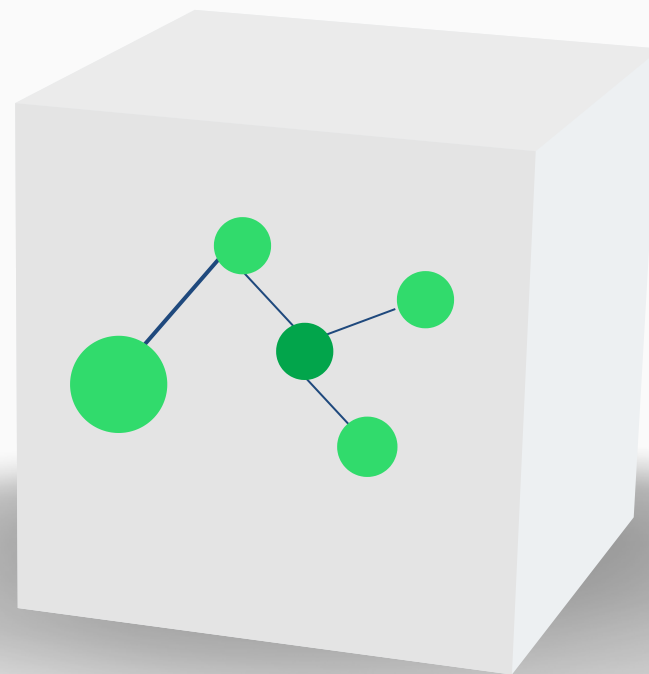
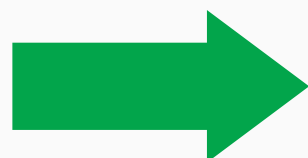
- RandomForest
- Multilayer-  
perception
- SMO

Mayor precisión

Mayor eficiencia

Menor interpretación  
y explicabilidad

Entradas



Salidas



Entradas



Salidas



# Objetivos

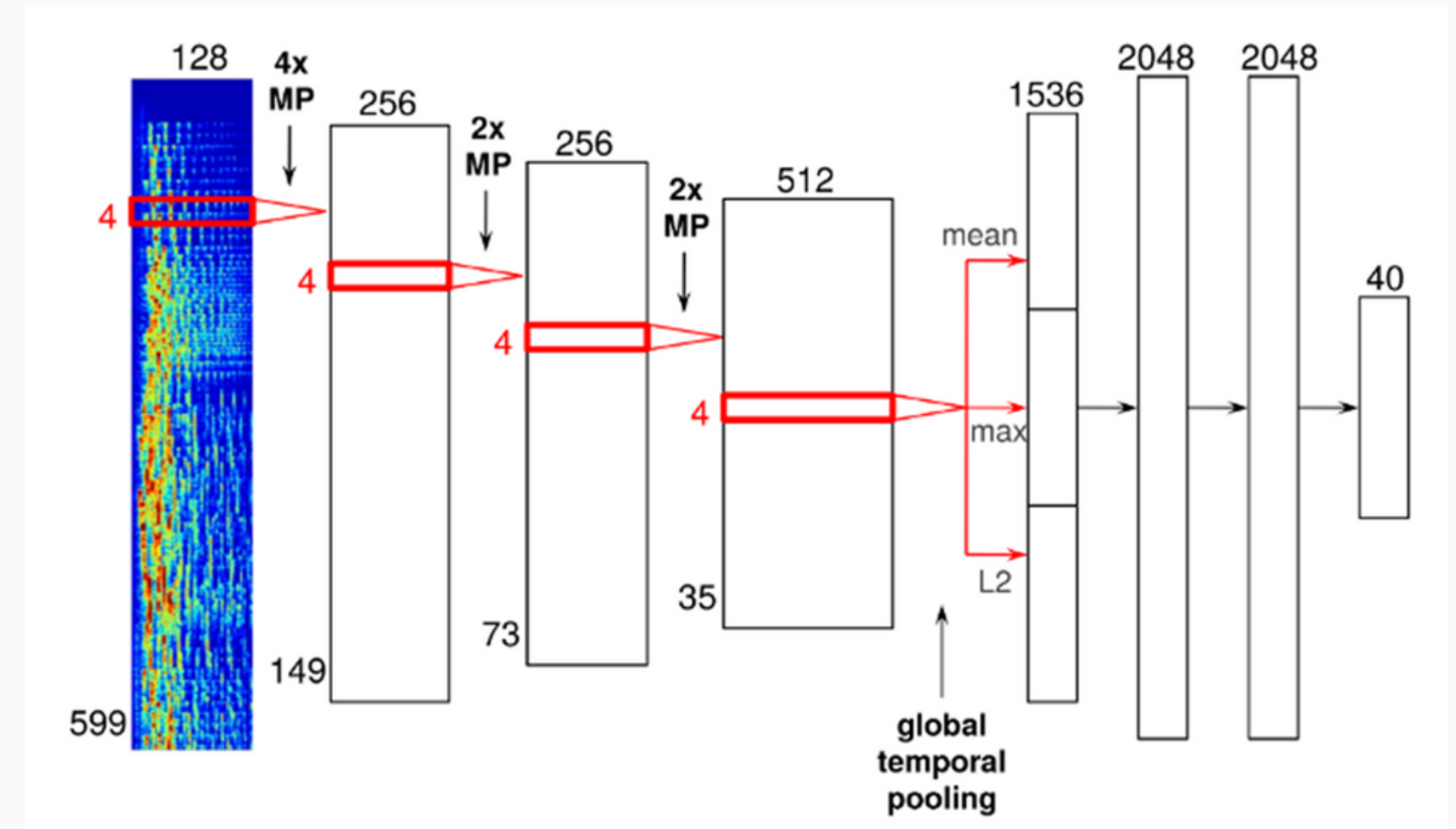
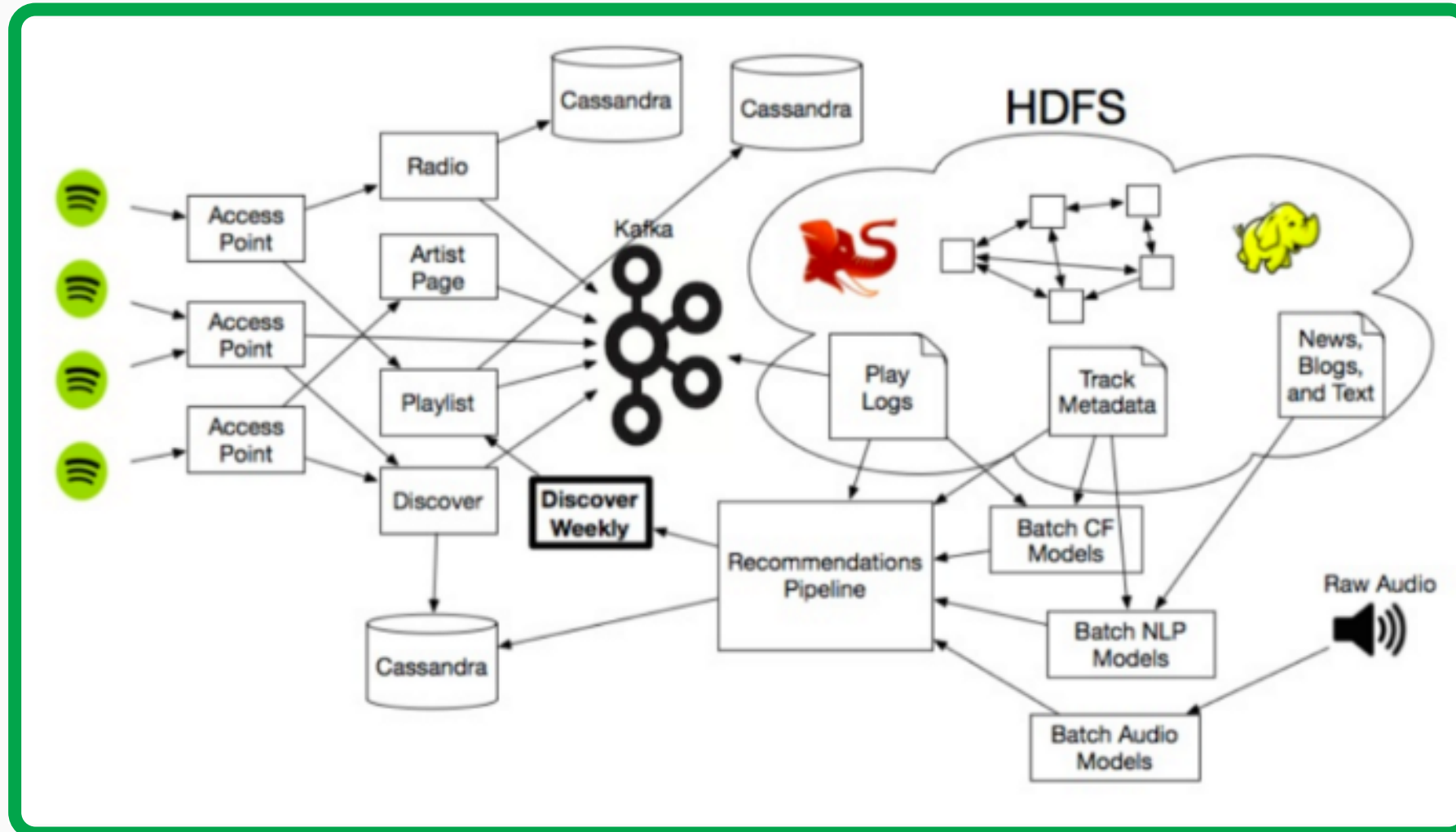
*¿Cuál de los algoritmos de recomendación usados más frecuentemente tiene una mayor precisión?  
¿Y cual una mayor explicabilidad?*

*¿Cuál de estos algoritmos daría el mejor resultado combinando precisión y explicabilidad?*

*¿En este tipo de problemas y cuestiones que convendría anteponer, una mayor precisión, o una mejor explicabilidad?*



# ¿Cómo funcionan los algoritmos de Spotify?



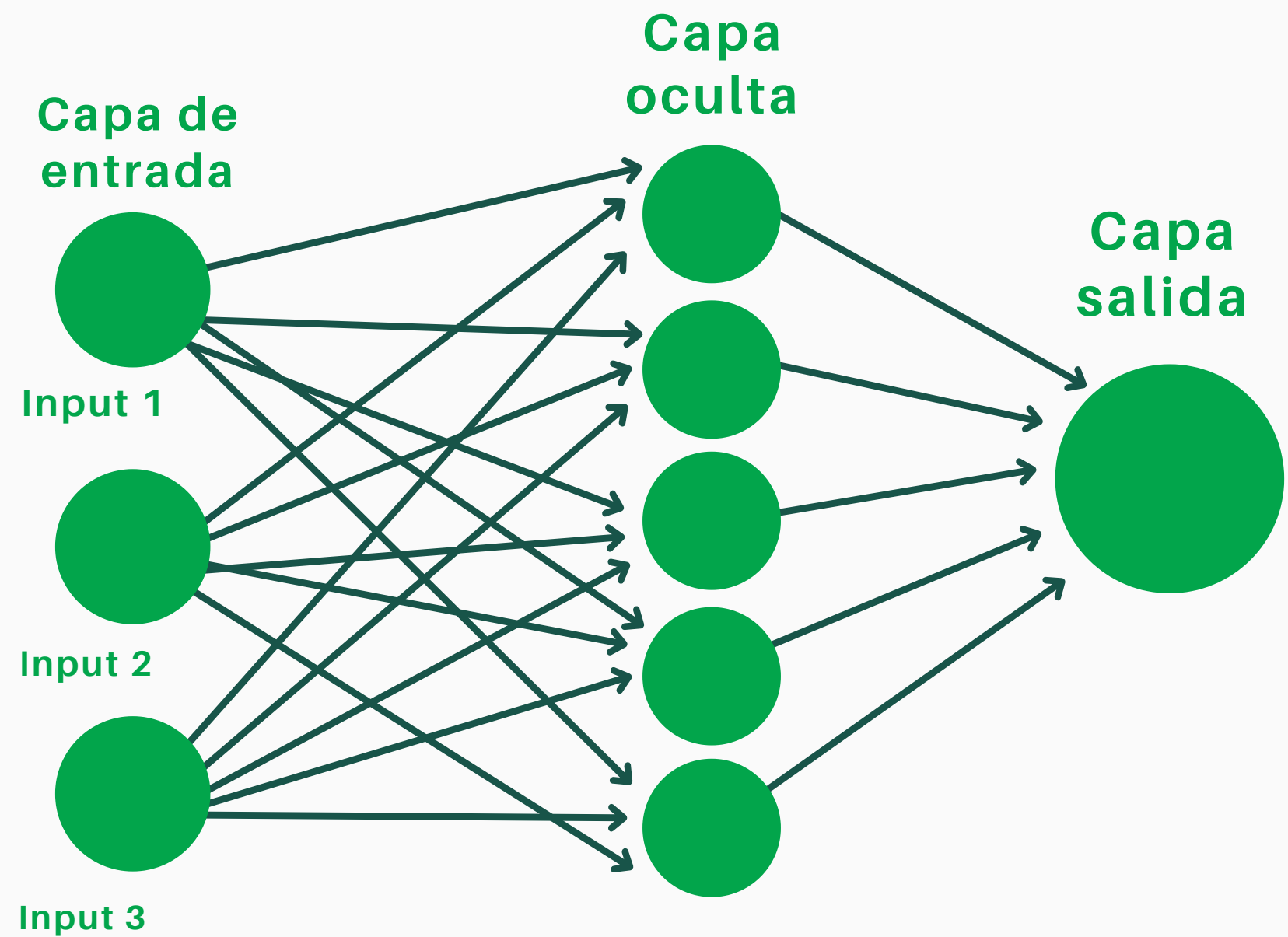
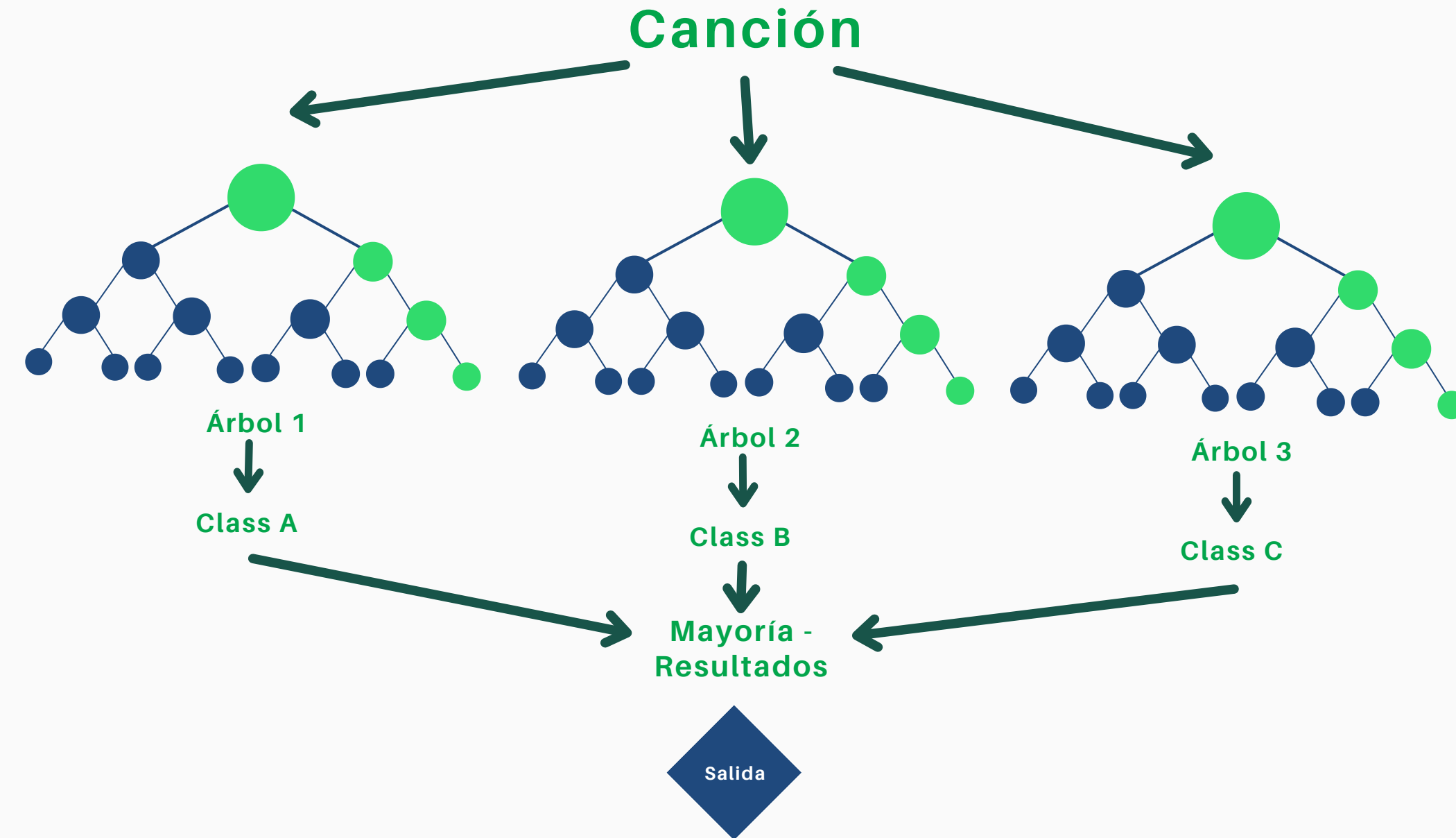
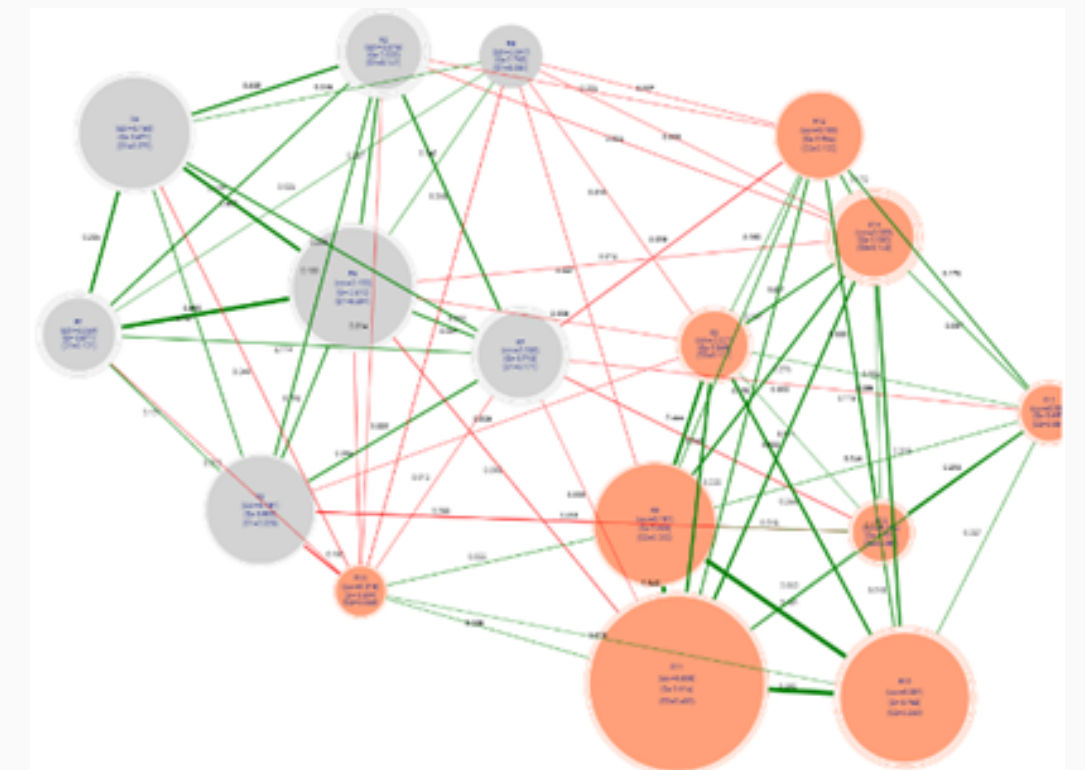
**Filtrado Colaborativo**  
Análisis del comportamiento

**Audio sin procesar**  
Análisis de las pistas de audio en bruto

**Procesamiento del Lenguaje Natural**  
Análisis de texto

# Cajas Negras

- *RandomForest*
- *Multilayer-perception*
- *SMO*



# Datos

	A	B	C	D	E	F	G	H	I	J	K	L	
1	id,acousticness,danceability,duration_ms,energy,instrumentalness,key,liveness,loudness,mode,speechiness,tempo,time_signature,valence,target,song_title,artist												
2	0,0.0102,0.833,204600,0.434,0.0219,2,0.165,-8.795,1,0.431,150.062,4,0,0.286,1,Mask Off,Future												
3	1,0.199,0.743,326933,0.359,0.00611,1,0.137,-10.401,1,0.0794,160.083,4,0,0.588,1,Redbone,Childish Gambino												
4	2,0.0344,0.838,185707,0.412,0.000234,2,0.159,-7.148,1,0.289,75.044,4,0,0.173,1,Xanny Family,Future												
5	3,0.604,0.494,199413,0.338,0.515,0.0922,-15.236,1,0.0261,86.468,4,0,0.23,1,Master Of None,Beach House												
6	4,0.18,0.678,392893,0.561,0.512,5,0.439,-11.648,0,0.0694,174.004,4,0,0.904,1,Parallel Lines,Junior Boys												
7	5,0.00479,0.804,251333,0.56,0,0,8,0.164,-6.682,1,0.185,85.023,4,0,0.264,1,Sneakinâ€™™,Drake												
8	6,0.0145,0.739,241400,0.472,7.27e-06,1,0.207,-11.204,1,0.156,80.03,4,0,0.308,1,Childs Play,Drake												
9	7,0.0202,0.266,349667,0.348,0.664,10,0.16,-11.609,0,0.0371,144.154,4,0,0.393,1,GyÃ¶ngyhajÃ© lÃ¡ny,Omega												
10	8,0.0481,0.603,202853,0.944,0,0,11,0.342,-3.626,0,0.347,130.035,4,0,0.398,1,I've Seen Footage,Death Grips												
11	9,0.00208,0.836,226840,0.603,0,0,7,0.571,-7.792,1,0.237,99.994,4,0,0.386,1,Digital Animal,Honey Claws												
12	10,0.0572,0.525,358187,0.855,0.0143,5,0.649,-7.372,0,0.0548,111.951,3,0,0.524,1,Subways - In Flagranti Extended Edit,The Avalanches												
13	11,0.0915,0.753,324880,0.748,0.00348,10,0.212,-8.62,1,0.0494,104.322,4,0,0.642,1,Donme Dolap - Baris K Edit,Modern Folk ÃœÃ§lÃ¼sÃ¼												
14	12,0.253,0.603,356973,0.434,0.0619,0,0.108,-11.062,1,0.0342,127.681,4,0,0.381,1,Cemalim,Erkin Koray												
15	13,0.366,0.762,243270,0.476,0,0,0,0.103,-12.686,1,0.114,130.007,4,0,0.367,1,One Night,Lil Yachty												
16	14,0.44,0.662,247288,0.603,0,0,9,0.0972,-8.317,0,0.0793,125.011,4,0,0.351,1,Oh lala,PNL												
17	15,0.019,0.637,188333,0.832,0.0563,6,0.316,-6.637,1,0.163,99.988,4,0,0.317,1,Char,Crystal Castles												
18	16,0.0239,0.603,270827,0.955,0.0451,1,0.119,-4.111,1,0.0458,123.922,4,0,0.773,1,World In Motion,New Order												
19	17,0.222,0.789,447007,0.659,0.00049,4,0.184,-12.654,0,0.0479,122.415,4,0,0.842,1,One Night Under a Green Eucalyptus												

Datos  
identificadores

Nombre de la  
canción, artista...

Variables  
del audio

acústica, duración,  
energía,  
sonoridad...

Opinión  
del usuario

Me gusta - 1  
No me gusta - 0

2000 canciones

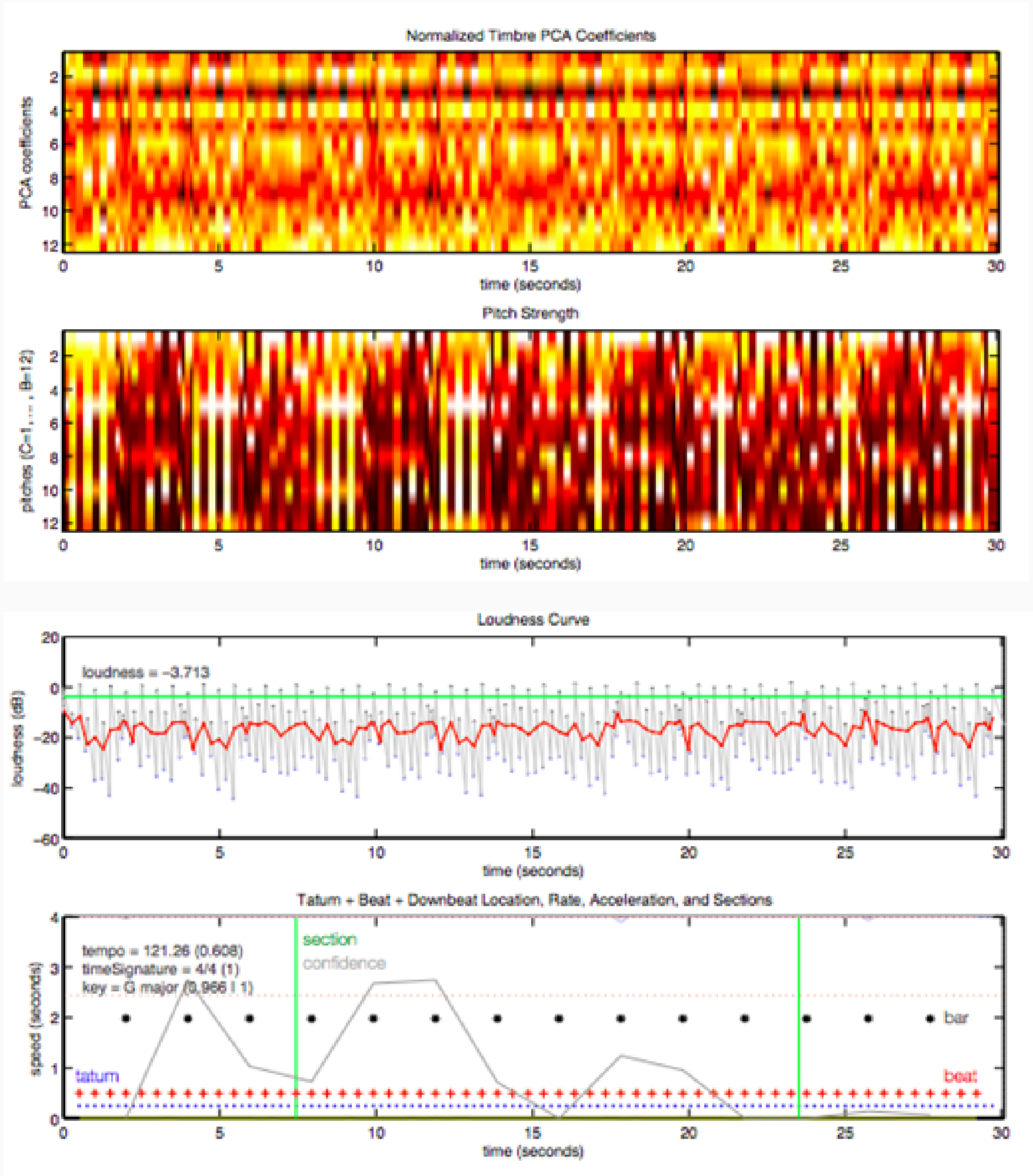


1000 gustan



1000 no gustan





Nombre	Definición
<i>Acousticness</i> (acústica)	Una medida de confianza de 0,0 a 1,0 de si la pista es acústica. 1,0 representa una alta confianza en que la pista es acústica. [0,1]
<i>Danceability</i> (bailabilidad)	La bailabilidad describe lo adecuada que es una pista para bailar basándose en una combinación de elementos musicales que incluyen el tempo, la estabilidad del ritmo, la fuerza del compás y la regularidad general. Un valor de 0.0 es el menos bailable y 1.0 es el más bailable.
<i>Duration_ms</i> (duración)	La duración de la pista en milisegundos.
<i>Energy</i> (energía)	La energía es una medida de 0,0 a 1,0 y representa una medida perceptiva de intensidad y actividad. Típicamente, las pistas energéticas se sienten rápidas, fuertes y ruidosas. Por ejemplo, el death metal tiene mucha energía, mientras que un preludio de Bach tiene una puntuación baja en la escala. Entre las características perceptivas que contribuyen a este atributo se encuentran el rango dinámico, el volumen percibido, el timbre, la velocidad de aparición y la entropía general.
<i>Key</i> (clave)	La tonalidad de la pista. Los números enteros se asignan a los tonos utilizando la notación estándar de Pitch Class. Por ejemplo, 0 = C, 1 = C#/Db, 2 = D, etc. Si no se detecta ninguna clave, el valor es [-1, 1]
<i>Liveness</i> (vivacidad)	Detecta la presencia de público en la grabación. Los valores de <i>liveness</i> más altos representan una mayor probabilidad de que la pista haya sido interpretada en directo. Un valor superior a 0,8 proporciona una fuerte probabilidad de que la pista sea en vivo.
<i>loudness</i> (sonoridad) <i>mode</i> (modo)	La sonoridad general de una pista en decibelios (dB). Los valores de sonoridad se promedian en toda la pista y son útiles para comparar la sonoridad relativa de las pistas. La sonoridad es la cualidad de un sonido que es el principal correlato psicológico de la fuerza física (amplitud). Los valores suelen os El modo indica la modalidad (mayor o menor) de una pista, el tipo de escala del que se deriva su contenido melódico. La mayor se representa con 1 y la menor con 0.
<i>speechiness</i>	La locuacidad detecta la presencia de palabras habladas en una pista. Cuanto más exclusivamente hablada sea la grabación (por ejemplo, un programa de entrevistas, un audiolibro o una poesía), más se acercará a 1,0 el valor del atributo. Los valores superiores a 0,66 describen pistas que probablemente estén compuestas exclusivamente por palabras habladas. Los valores entre 0,33 y 0,66 describen pistas que pueden contener tanto música como voz, ya sea en secciones o en capas, incluyendo casos como la música rap. Los valores por debajo de 0,33 representan probablemente música y otras pistas no habladas.
<i>tempo</i>	El tempo global estimado de una pista en pulsaciones por minuto (BPM). En la terminología musical, el tempo es la velocidad o el ritmo de una pieza determinada y se deriva directamente de la duración media de los tiempos.
<i>time_signature</i> (tiempo musical)	Una firma de tiempo estimada. La signatura de tiempo (metro) es una convención de notación para especificar cuántos tiempos hay en cada compás (o medida). La signatura de tiempo va de 3 a 7, indicando signaturas de tiempo de "3/4", a "7/4". Rango de valores: [3,7]
<i>valence</i> (valencia)	Una medida de 0,0 a 1,0 que describe la positividad musical que transmite una pista. Las pistas con alta valencia suenan más positivas (por ejemplo, felices, alegres, eufóricas), mientras que las pistas con baja valencia suenan más negativas (por ejemplo, tristes, deprimidas, enfadadas).Rango de valores: [0,1]
<i>Instrumentalness</i> (instrumentalidad)	Predice si una pista no contiene voces. Los sonidos "Ooh" y "aah" se tratan como instrumentales en este contexto. Las pistas de rap o de palabras habladas son claramente "vocales". Cuanto más se acerque el valor de instrumentalización a 1,0, mayor será la probabilidad de que la pista no tenga contenido vocal. Los valores superiores a 0,5 representan pistas instrumentales, pero la confianza es mayor a medida que el valor se acerca a 1,0.

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Classifier

Choose

J48 -C 0.25 -M 2

Test options

Use training set

Supplied test set

Cross-validation

Percentage split

Set...

Folds

10

%

66

More options...

(Nom) class

Start

Stop

Result list (right-click for options)

20:21:20 - trees.J48

Status

OK

Classifier output

=== Classifier model (full training set) ===

J48 pruned tree

-----

energy <= 0.177: 1 (74.0/3.0)

energy > 0.177

instrumentalness <= 0.0013

speechiness <= 0.124

danceability <= 0.779

duration\_ms <= 163213

mode <= 0

duration\_ms <= 157001: 2 (11.0/1.0)

duration\_ms > 157001: 1 (2.0)

mode > 0

instrumentalness <= 0.000016

time\_signature <= 3: 1 (2.0/1.0)

time\_signature > 3

key <= 1: 1 (4.0)

key > 1

acousticness <= 0.0119: 1 (5.0/1.0)

acousticness > 0.0119: 2 (18.0/6.0)

instrumentalness > 0.000016: 2 (7.0)

duration\_ms > 163213

TP Rate	FP Rate	Precision	Recall	F-Measure
0,716	0,260	0,729	0,716	0,723
0,740	0,284	0,727	0,740	0,734
0,728	0,272	0,728	0,728	0,728

TP Rate	FP Rate	Precision	Recall	F-Measure
0,694	0,309	0,687	0,694	0,691
0,691	0,306	0,698	0,691	0,695
0,693	0,307	0,693	0,693	0,693

TP Rate	FP Rate	Precision	Recall	F-Measure
0,711	0,293	0,703	0,711	0,707
0,707	0,289	0,715	0,707	0,711
0,709	0,291	0,709	0,709	0,709

TP Rate	FP Rate	Precision	Recall	F-Measure
0,679	0,258	0,720	0,679	0,699
0,742	0,321	0,703	0,742	0,722
0,711	0,290	0,711	0,711	0,711

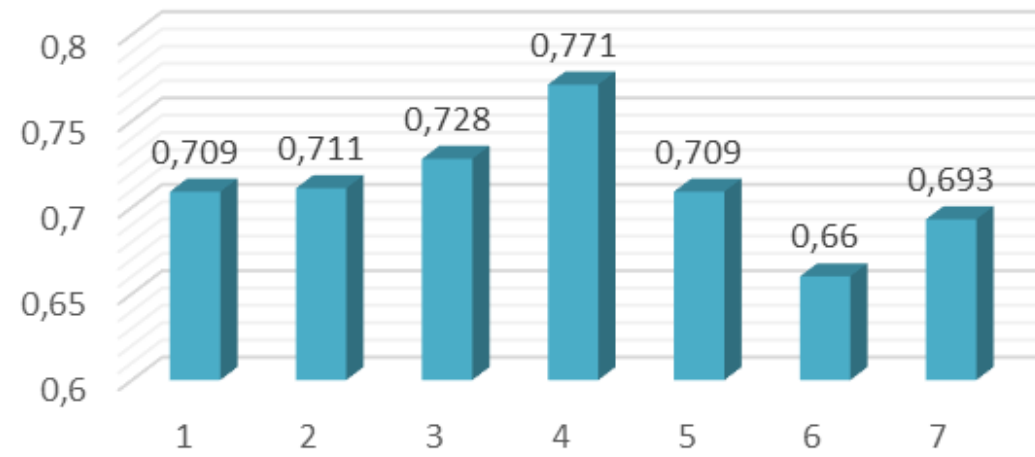
TP Rate	FP Rate	Precision	Recall	F-Measure
0,734	0,410	0,637	0,734	0,682
0,590	0,266	0,694	0,590	0,638
0,661	0,337	0,666	0,661	0,660

TP Rate	FP Rate	Precision	Recall	F-Measure
0,777	0,234	0,764	0,777	0,771
0,766	0,223	0,779	0,766	0,772
0,771	0,228	0,772	0,771	0,771

TP Rate	FP Rate	Precision	Recall	F-Measure
0,733	0,315	0,695	0,733	0,714
0,685	0,267	0,724	0,685	0,704
0,709	0,290	0,710	0,709	0,709

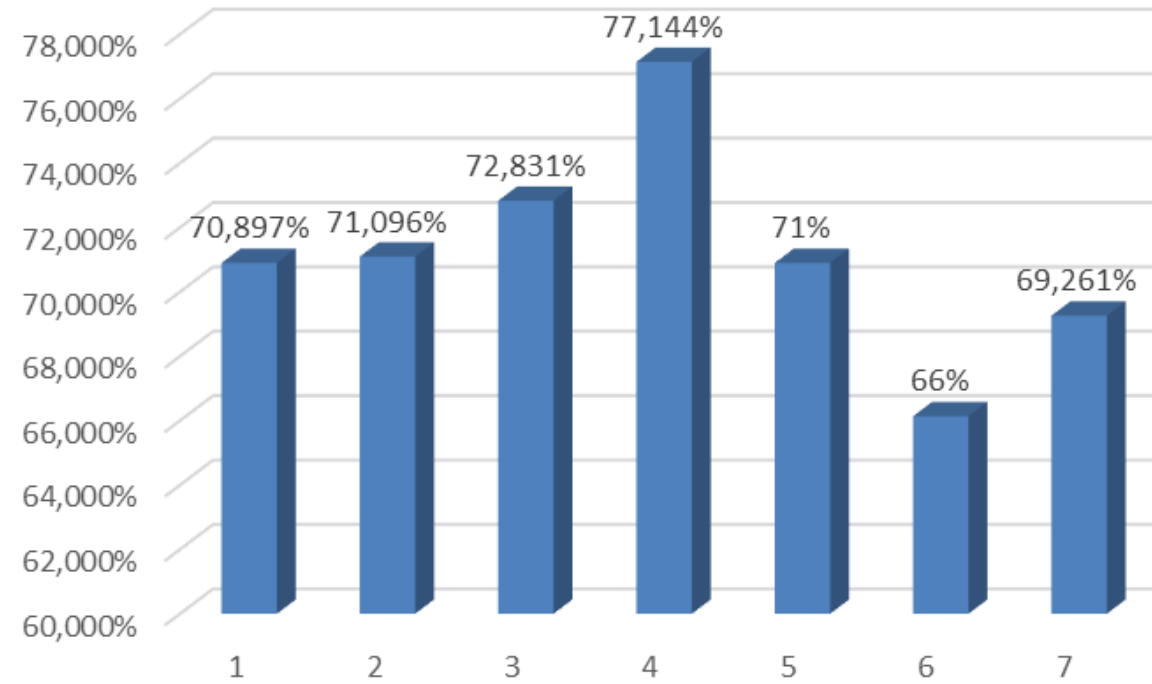
## F-Measure

Una medida combinada de precisión y recuperación calculada como  $2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$



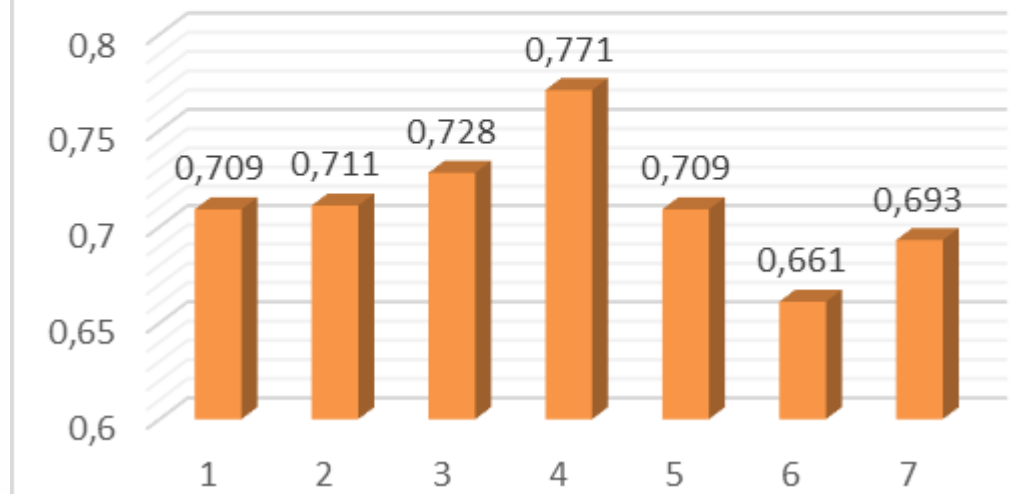
## Correct Cl. (%)

El porcentaje de instancias clasificadas correctamente suele denominarse precisión o exactitud de la muestra. Destaca RandomForest.



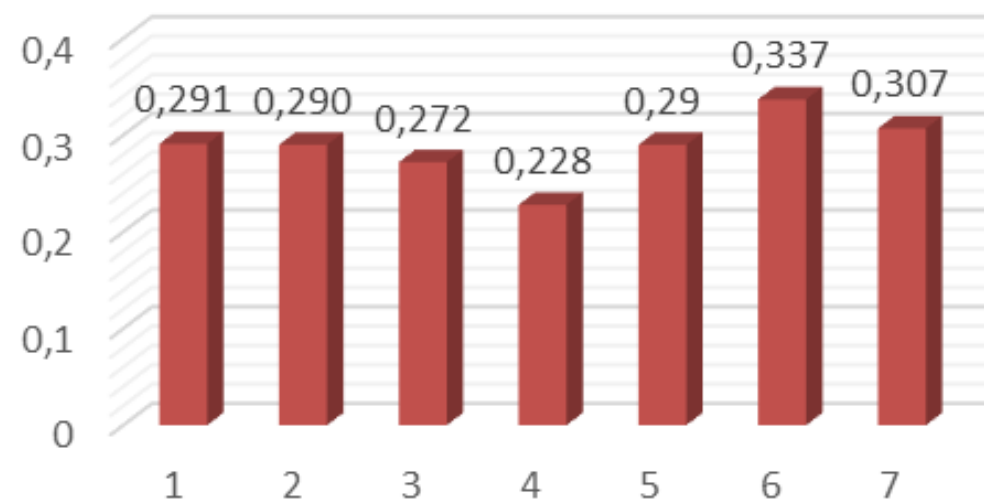
## Recall

Proporción de instancias clasificadas como una clase determinada dividida por el total real de esa clase (equivalente a la tasa TP)

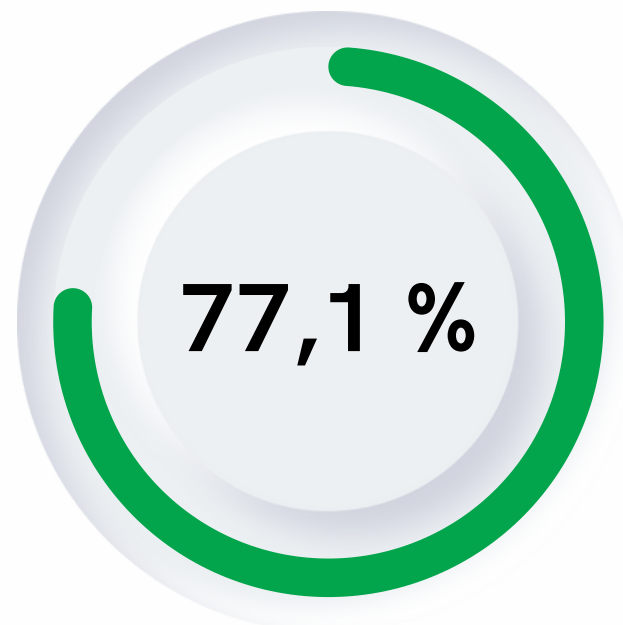


## FP Rate

Tasa de falsos positivos (instancias falsamente clasificadas como una clase determinada)

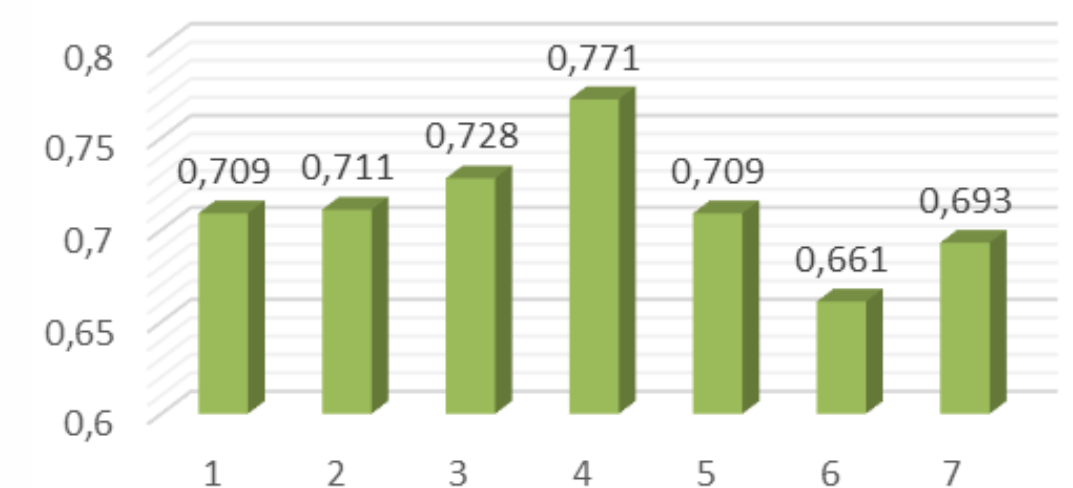


## Random Forest



## TP Rate (0-1)

Tasa de verdaderos positivos (instancias correctamente clasificadas como una clase determinada)





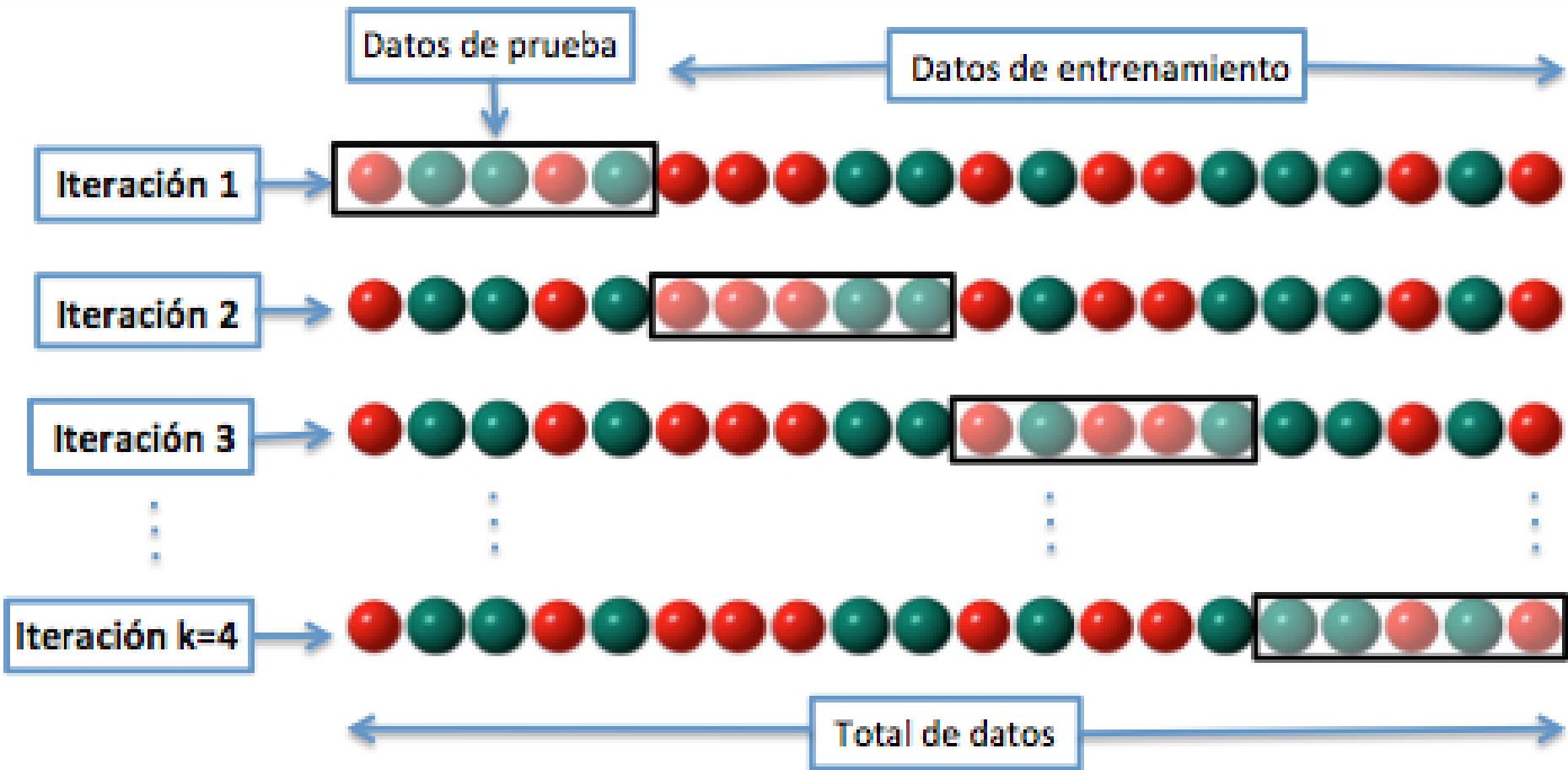
# Matriz de confusión

=== Confusion Matrix ===

```
a    b    <-- classified as
709 288 |    a = 1
299 721 |    b = 2
```

	N (modelo)	S (modelo)
n (real)	Negativos reales	Falsos Positivos
p (real)	Falsos Negativos	Positivos Reales

# Validación cruzada



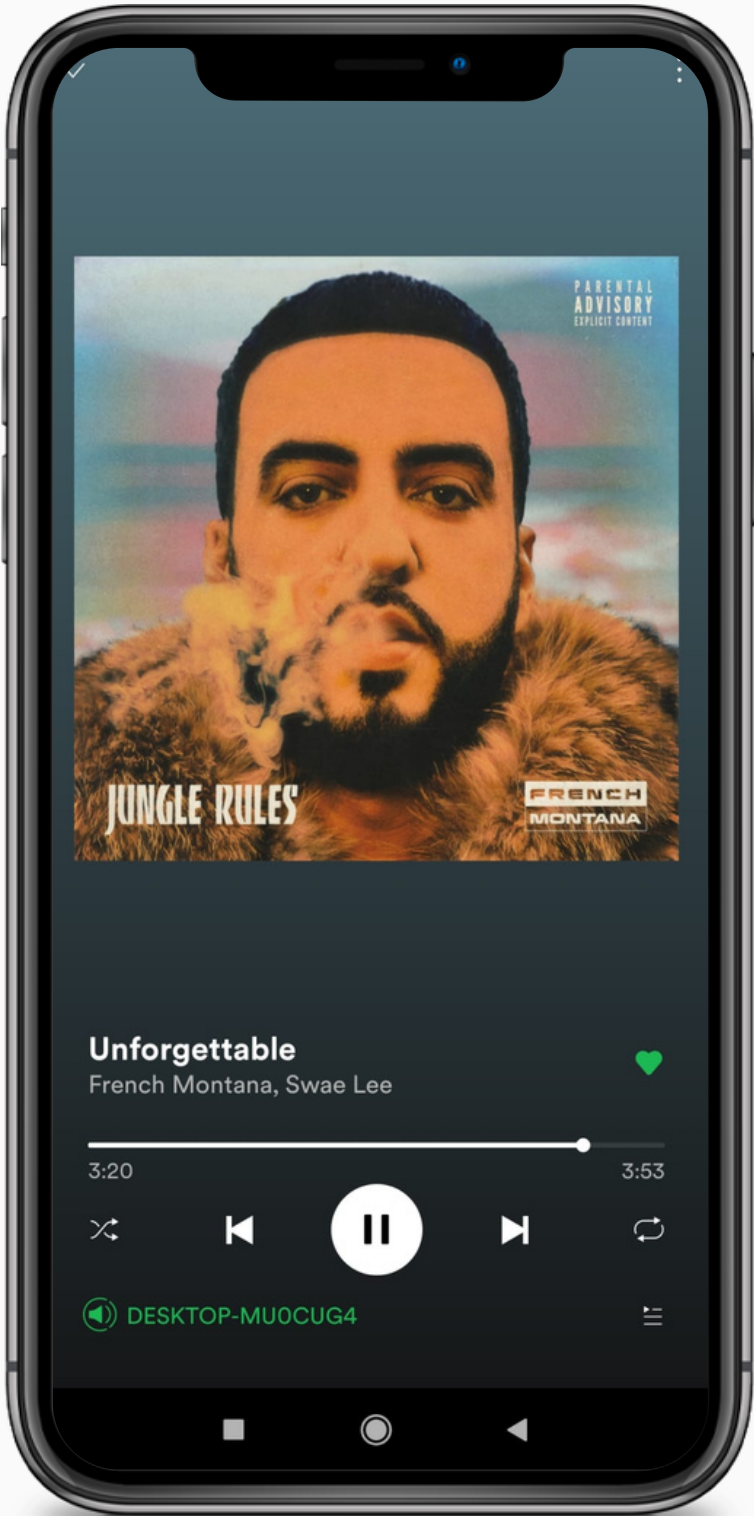
# Canción que suele gustar



Name	Value	Property name	Property value
acousticness	0.0293	Low	0.33 , 0
		Medium	0.66 , 0.33
		High	0.99 , 0.66
danceability	0.726	Low	0.41 , 0.12
		Medium	0.7 , 0.41
		High	0.98 , 0.7
duration_ms	233833	Low	345570.33 , 16042
		Medium	675098.67 , 345570.33
		High	1004627 , 675098.67
energy	0.769	Low	0.34 , 0.01
		Medium	0.67 , 0.34
		High	1 , 0.67
instrumentalness	0.0101	Low	0.33 , 0
		Medium	0.65 , 0.33
		High	0.98 , 0.65
key	6	Low	3.67 , 0
		Medium	7.33 , 3.67
		High	11 , 7.33
liveness	0.104	Low	0.34 , 0.02
		Medium	0.65 , 0.34
		High	0.97 , 0.65
loudness	-5.043	Low	-22.17 , -33.1
		Medium	-11.24 , -22.17
		High	-0.31 , -11.24

# Canción que no suele gustar

Name	Value	Property name	Property value
acousticness	0.0426	Low	0.33 , 0
		Medium	0.66 , 0.33
		High	0.99 , 0.66
danceability	0.794	Low	0.41 , 0.12
		Medium	0.7 , 0.41
		High	0.98 , 0.7
duration_ms	248413	Low	345570.33 , 16042
		Medium	675098.67 , 345570.33
		High	1004627 , 675098.67
energy	0.811	Low	0.34 , 0.01
		Medium	0.67 , 0.34
		High	1 , 0.67
instrumentalness	0.000001	Low	0.33 , 0
		Medium	0.65 , 0.33
		High	0.98 , 0.65
key	6	Low	3.67 , 0
		Medium	7.33 , 3.67
		High	11 , 7.33
liveness	0.101	Low	0.34 , 0.02
		Medium	0.65 , 0.34
		High	0.97 , 0.65
loudness	-8.966	Low	-22.17 , -33.1
		Medium	-11.24 , -22.17
		High	-0.31 , -11.24





# Lenguaje natural

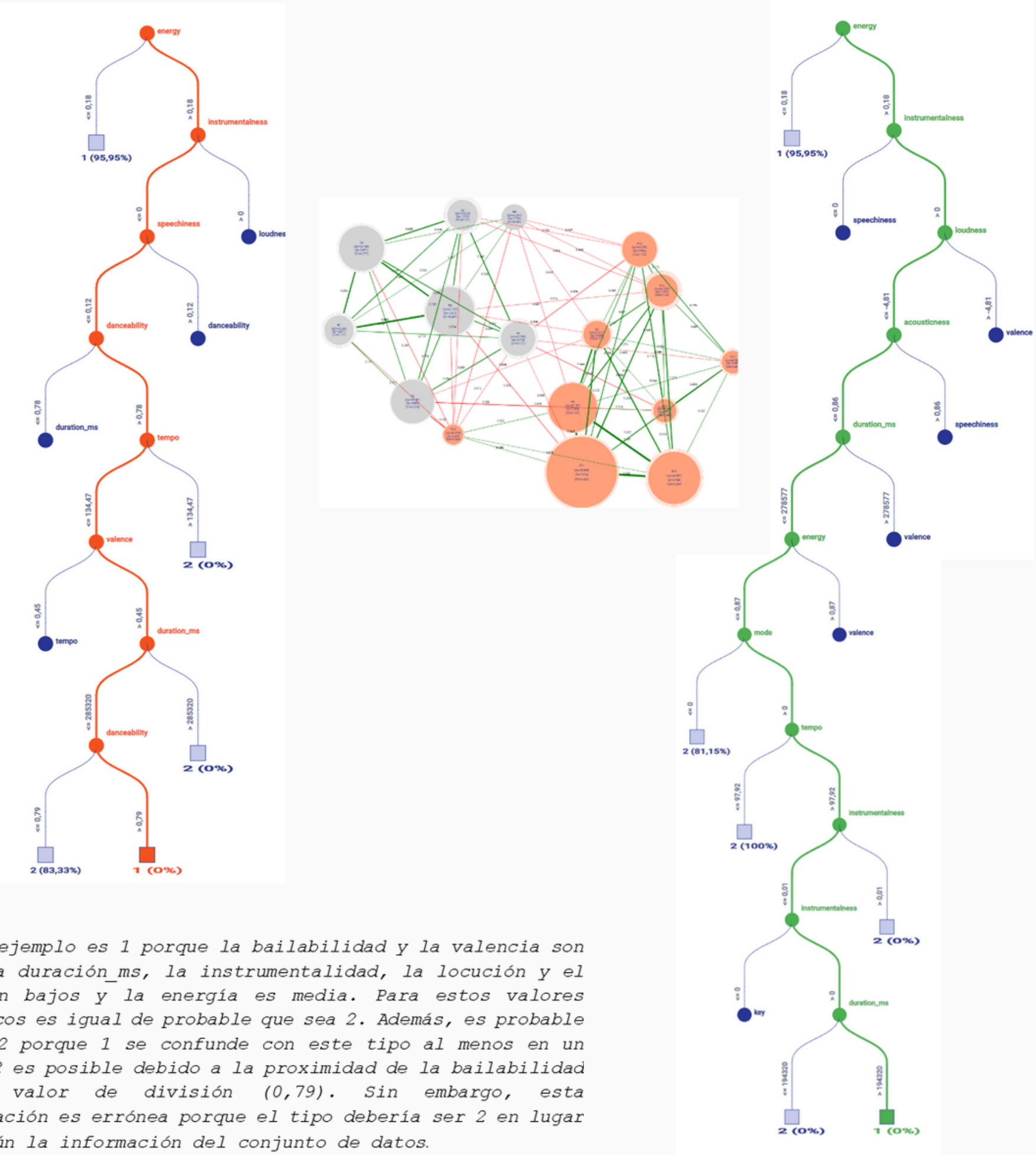
Algorithm	Local explanation
J48	Example is 1 because acousticness, duration_ms, instrumentalness, loudness and mode are low and energy and tempo are medium.
FURIA	We have a high confidence in the classification result. It is very likely that this example is 2 because loudness and liveness are low and instrumentalness and energy are medium. However, this is wrong because the type should be 1 instead of 2 according to the information in the dataset.
RandomForest	This instance is classified as 1

Algorithm	Local explanation
J48	Example is 1 because danceability and valence are high, duration_ms, instrumentalness, speechiness and tempo are low and energy is medium. However, this classification is wrong because type should be 2 instead of 1 according to the information in the dataset.
FURIA	Classifier performs stretching to determine that example is 1. However, this is wrong because the type should be 2 instead of 1 according to the information in the dataset.
RandomForest	This instance is classified as 2

El ejemplo es 2 porque la bailabilidad, el tempo y la valencia son altos, la acústica, la instrumentalidad y la sonoridad son bajos y la duración\_ms, la vivacidad y el discurso son medios. Además, es probable que sea 1 porque 2 se confunde con este tipo al menos en un 10%. Pero es poco probable que sea 1. El 1 es posible debido a la proximidad del tempo con el valor de división (111,0).

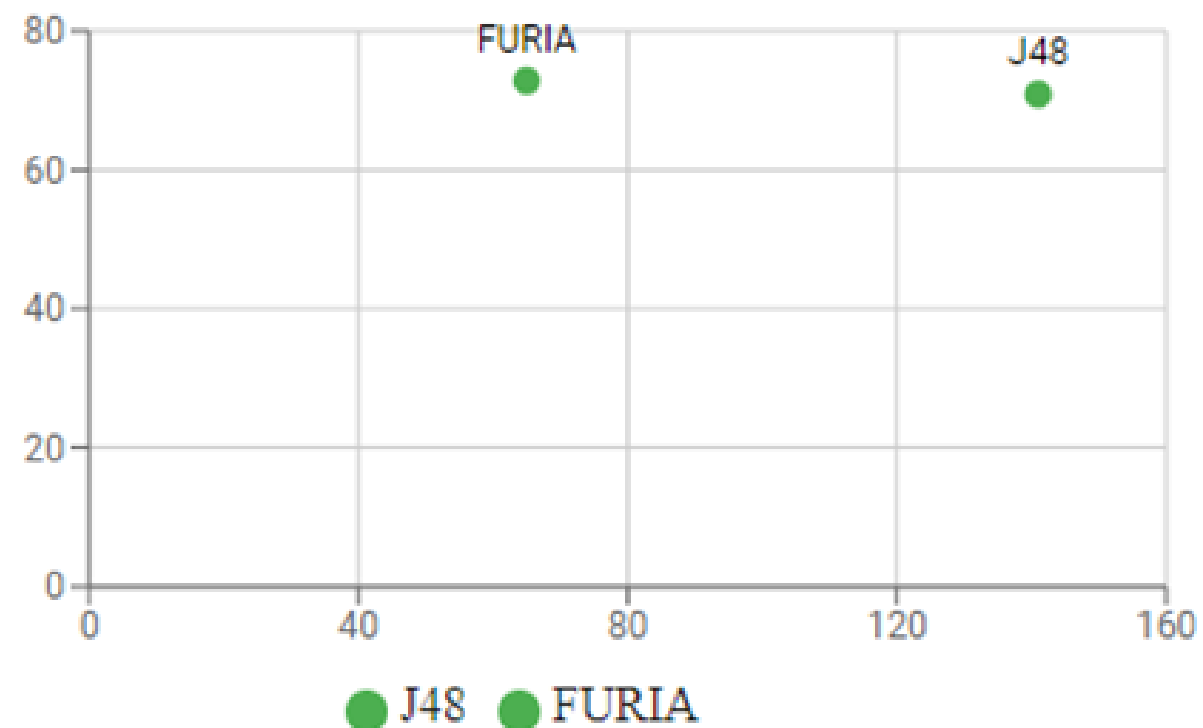
El clasificador realiza un estiramiento para determinar que el ejemplo es 1. Sin embargo, esto es incorrecto porque el tipo debería ser 2 en lugar de 1 según la información del conjunto de datos.

El ejemplo es 1 porque la bailabilidad y la valencia son altas, la duración\_ms, la instrumentalidad, la locución y el tempo son bajos y la energía es media. Para estos valores específicos es igual de probable que sea 2. Además, es probable que sea 2 porque 1 se confunde con este tipo al menos en un 10%. El 2 es posible debido a la proximidad de la bailabilidad con el valor de división (0,79). Sin embargo, esta clasificación es errónea porque el tipo debería ser 2 en lugar de 1 según la información del conjunto de datos.

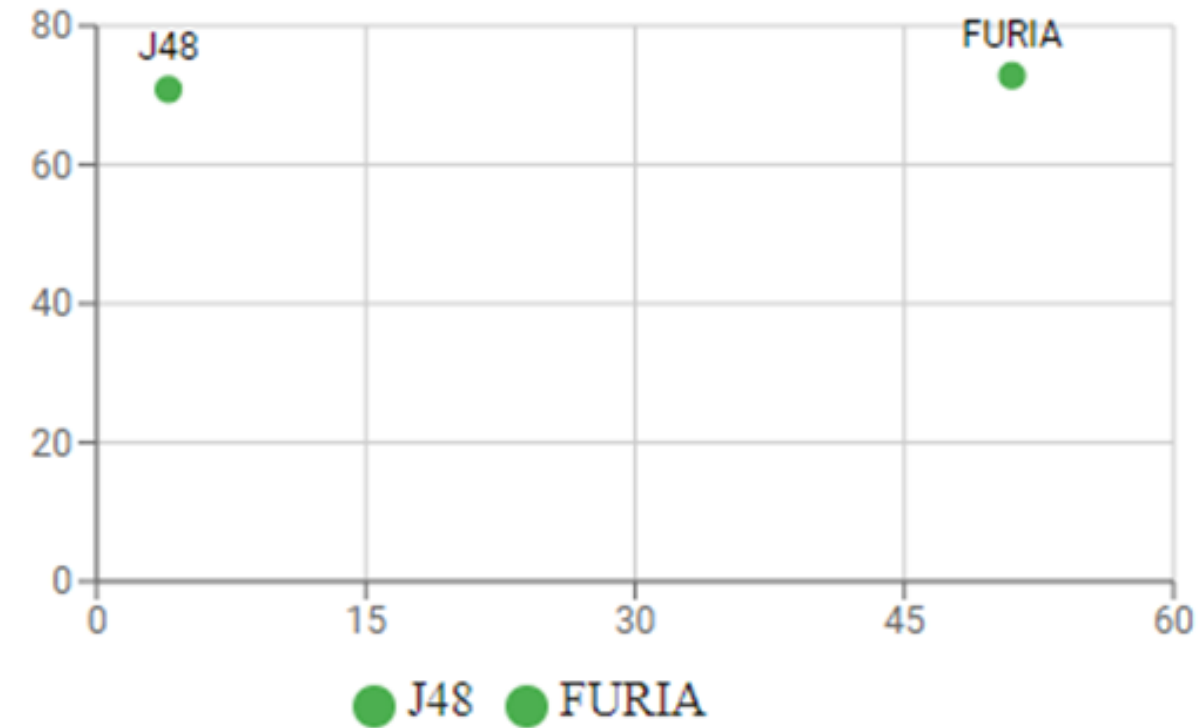


# Precisión vs. Explicabilidad

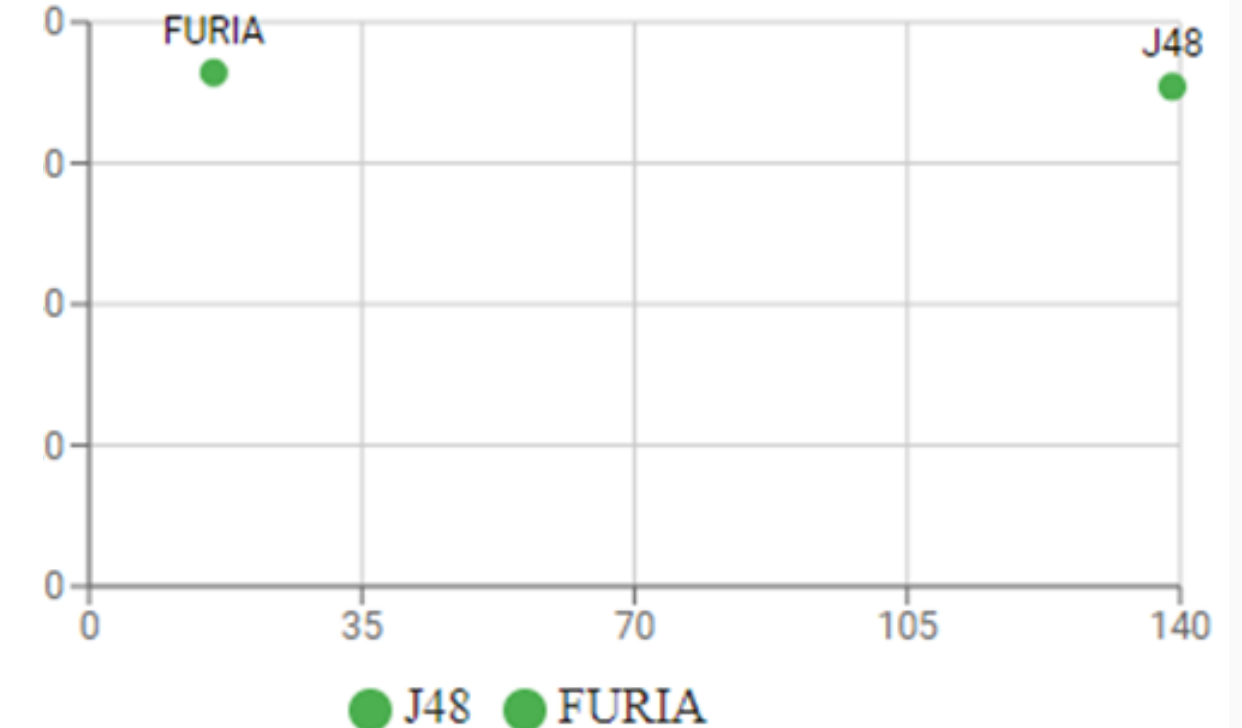
Precision (CCI) vs Interpretabilidad (TRL)



Precision (CCI) vs Interpretabilidad (nº de conceptos)



Precision (CCI) vs Interpretabilidad (nº de reglas)



**Menor número de  
conceptos:  
J48**

**Menor número de reglas  
obtiene mayor CCI:  
FURIA**

# Conclusiones

*El algoritmo que presenta una mejor resolución en las métricas de precisión es RandomForest con un 77,14% de efectividad.*

*Algoritmos como J48, tienen una diferencia de precisión no muy notable frente a RandomForest, y desarrolla una mejor explicación de la resolución dada en los ejemplos utilizados.*

*La solución para este problema sería confiar de la explicación de las cajas blancas J48 y FURIA, cuando hay consenso con las cajas negras como RandomForest, o cuando los 3 coincidan.*