



Una propuesta para la clasificación emocional de un álbum a partir de la letra de sus canciones

A Proposal for the Emotional Classification of an Album Using the Lyrics of its Songs

Francisco Javier Moreno Arboleda¹, John Freddy Duitama Muñoz²,
Luis Fernando Montoya Gómez³

Fecha de recepción: 10 de noviembre de 2014

Fecha de aceptación: 18 de septiembre de 2015

Cómo citar: Francisco Javier Moreno Arboleda, John Freddy Duitama Muñoz, & Luis Fernando Montoya Gómez. (2016). Una propuesta para la clasificación emocional de un álbum a partir de la letra de sus canciones. *Revista Tecnura*, 20(47), 57-70. doi: <http://dx.doi.org/10.14483/udistrital.jour.tecnura.2016.1.a05>

Resumen

Los grandes volúmenes de datos que se manejan actualmente demandan métodos automáticos para la extracción de conocimiento. En particular, la minería de texto se ocupa de la extracción de conocimiento a partir de textos. En este artículo se proponen dos algoritmos para determinar a partir de las letras de sus canciones, la tendencia emocional de un álbum. Se sigue un enfoque jerárquico: las categorías emocionales para clasificar los álbumes agrupan a las subcategorías emocionales de las canciones. Esto es razonable, porque una canción tiende a estar orientada a una (sub)categoría emocional específica. De esta forma la categoría emocional de un álbum es una ponderación de las subcategorías emocionales de sus canciones. Esta ponderación puede ser configurada por parte del analista musical, lo que permite incorporar un elemento de subjetividad en la propuesta. Mediante una serie de experimentos se evaluaron los algoritmos propuestos. Aunque es

necesario experimentar con más datos, los resultados evidenciaron la conveniencia, consistencia y utilidad de los algoritmos propuestos.

Palabras clave: análisis musical, categorías emocionales, clasificación, minería de texto.

Abstract

Current volumes of data require automated methods for knowledge extraction. In particular, text mining deals with extracting knowledge from texts. In this paper we propose two algorithms to determine, from the lyrics of its songs, the emotional tendency of an album. We follow a hierarchical approach: the emotional categories to classify the albums group the emotional subcategories of the songs. This is reasonable because a song tends to be oriented to a specific emotional (sub) category. In this way, the emotional category of an album is a weighting of the emotional subcategories of its songs. This weighting can be customized by the musical analyst,

¹ Ingeniero de Sistemas, magister en Ingeniería de Sistemas, doctor en Ingeniería-Sistemas. Docente de la Universidad Nacional de Colombia. Medellín, Colombia. Contacto: ijmoreno@unal.edu.co

² Ingeniero de Sistemas, magister en Ingeniería de Sistemas, doctor en Informática. Docente de la Universidad de Antioquia. Medellín, Colombia. Contacto: john.duitama@udea.edu.co

³ Ingeniero de sistemas. Universidad Nacional de Colombia, Sede Medellín. Medellín, Colombia. Contacto: lfmontoyag@unal.edu.co

which allows incorporating a subjective element in our proposal. Through a series of experiments we evaluated our algorithms. Although it is necessary to experiment with more data, our results show the

expediency, consistency, and usefulness of our proposed algorithms.

Keywords: classification, emotional categories, musical analysis, text mining.

INTRODUCCIÓN

Entre las funcionalidades que ofrecen sitios de música como *Allmusic.com*, *Stereomood.com*, *Musicoverly.com*, *www.musicoverly.com* y *Friendllymusic.com*, entre otros, está la clasificación de las canciones en categorías emocionales. Esto permite a los usuarios crear listas de reproducción y encontrar canciones relacionadas con su estado de ánimo. En estos sitios la clasificación emocional de las canciones se logra a partir de la retroalimentación de los usuarios. Otra alternativa es obtener esta clasificación teniendo en cuenta las características inherentes de una canción como su letra y su sonido. Para ello se pueden aplicar métodos de *minería de texto* y *de audio*.

Por ejemplo, en Hu, Downie y Ehmann (2009) se presenta un método para clasificar canciones en 18 categorías emocionales. Para esto se consideraron características lingüísticas (la letra de la canción) y espectrales del audio. Para evaluar la precisión de la clasificación, la cual estuvo alrededor del 60 %, se aplicaron diferentes técnicas a la letra: *bag-of-words* (Ferragina y Scaiella, 2011), *part-of-speech* (Diesner, 2006), *function words* (Roberts, 2002) y *stemming* (Jivani, 2011). Además, los autores mostraron que la incorporación de las características espectrales del audio no mejoró significativamente la precisión de la clasificación.

En Panagakakis y Kotropoulos (2011) se propone el algoritmo LRRC (*low-rank representation-based Classification*) para la clasificación emocional de canciones, el cual solo usa las características espectrales del audio. Los experimentos se hicieron con 180 bandas sonoras de películas y seis categorías emocionales: *happiness* (felicidad), *sadness*

(tristeza), *fear* (temor), *anger* (ira), *surprise* (sorpresa) y *tenderness* (ternura). El LRRC se comparó con algunos de los algoritmos más usados en tareas de clasificación: *sparse representations-based classifier* (Zhang et al., 2012), *support vector machines* (SVM) (Fletcher, 2009; Chang y Lin, 2011) y *nearest neighbor* (Liu, Moore, Gray y Yang, 2004); el LRRC obtuvo la precisión más alta (alrededor del 64 %).

En Trohidis, Tsoumakas, Kalliris y Vlahavas (2011) se evalúa el rendimiento de cuatro algoritmos de clasificación emocionales de canciones: *label powerset* (Cherman, Monard y Metz, 2011), *binary relevance* (Read, Pfahringer, Holme y Frank, 2011), *Random k-labelsets* (Tsoumakas, Katakis y Vlahavas, 2011) y *multilabel k-nearest neighbor*. Los autores consideraron que una canción puede pertenecer simultáneamente a varias categorías emocionales. Se consideraron seis categorías emocionales: *amazement-surprise* (asombro-sorpresa), *happiness-satisfaction* (felicidad-satisfacción), *relaxation-calmness* (relajación-calma), *quietness-stillness* (tranquilidad-quietud), *sadness-loneliness* (tristeza-soledad) y *anger-fear* (enojo-temor). El algoritmo *random k-labelsets* obtuvo la precisión más alta (alrededor del 80 %).

En Li y Ogiyara (2013) se definieron 13 categorías emocionales para la clasificación de las canciones. Se consideraron 30 características espectrales obtenidas mediante MARSYAS, una herramienta para el procesamiento de audio orientada a las aplicaciones de recuperación de información de la música. Como método de clasificación se usó LIBSVM (Chang y Lin, 2011) (*A Library for Support Vector Machines*), el cual se fundamenta en los SVM. La precisión estuvo alrededor del 46 %.

En el presente artículo se presenta una propuesta para obtener la *clasificación emocional de un álbum* a partir de la clasificación emocional de sus canciones. La propuesta es flexible porque permite que el usuario especifique el grado con el que las categorías emocionales asociadas con las canciones contribuyen a la categoría emocional de un álbum; por ejemplo, permite la incorporación de elementos *subjetivos* en la ponderación de las emociones.

El artículo está organizado así. Primero, se presentan los fundamentos teóricos para lograr la clasificación emocional de una canción. Posteriormente se expone la propuesta para la clasificación emocional de un álbum. Luego se muestran los experimentos y el análisis de los resultados. Por último, se plantean algunas conclusiones y se sugieren trabajos futuros.

FUNDAMENTOS TEÓRICOS

Para ejemplificar las técnicas descritas en esta sección para analizar la letra de una canción y obtener su clasificación emocional, se trabajará con la letra de la canción "Come on, get happy" del grupo The Partridge Family. La letra de la canción es:

*Hello world, hear the song that we're singing
Come on, get happy
A whole lotta loving is what we'll be bringing
We'll make you happy
We had a dream we'd go traveling together
And spread a little loving, then we'll keep moving on
Something always happens whenever we're together
We get a happy feeling when we're singing a song
Traveling along, there's a song that we're singing
Come on, get happy
A whole lotta loving is what we'll be bringing
We'll make you happy
We'll make you happy
We'll make you happy*

Paso 1. Se convierte la letra de la canción a minúsculas, se eliminan los saltos de línea y todos los caracteres que no son letras. Nótese que al eliminar el apóstrofo, también se eliminan las letras correspondientes a las contracciones 're, 'll, 'd, 's.

hello world hear the song that we singing come on get happy a whole lotta loving is what we be bringing we make you happy we had a dream we go traveling together and spread a little loving then we keep moving on something always happens whenever we together we get a happy feeling when we singing a song traveling along there a song that we singing come on get happy a whole lotta loving is what we be bringing we make you happy we make you happy we make you happy.

Paso 2. Se eliminan las *stop words* (Feldman y Sanger, 2007), es decir, aquellas palabras que aportan poco o nada de significado para clasificar emocionalmente el texto, como artículos, pronombres o preposiciones:

hello world hear song singing come get happy lotta loving bringing make happy dream go traveling spread loving keep moving happens get happy feeling singing song traveling song singing come get happy lotta loving bringing make happy make happy make happy

Paso 3. Se aplica la técnica de *stemming*, de la cual se obtiene la raíz de cada palabra. Por ejemplo, la raíz de la palabra *feeling* es *feel*. Existen varios algoritmos para realizar el *stemming*, entre los más conocidos están el de Porter (Noraida y Noor, 2012), el de Snowball (Agichtein y Gravano, 2000) y el de Lovins (Jivanis, 2011). En el presente artículo se usa el algoritmo de Snowball, ya que con este se obtiene la mejor precisión y es ampliamente difundido.

hello world hear song sing come get happi lotta love bring make happi dream go travel spread love keep move happen get happi feel sing song travel song sing come get happi lotta love bring make happi make happi make happi

Nótese que el algoritmo de Snowball genera la palabra *happi* como la raíz de la palabra *happy*, debido a que una de sus reglas es $(*v*) Y \rightarrow I$ (es decir que, al tener una vocal entre consonantes seguida de la letra *Y*, se cambia la *Y* por *I*) (Sharma, 2012).

Paso 4. Los SVM (*support-vector machines*) son un tipo de algoritmos de aprendizaje supervisado, desarrollado por Vapnik (Fletcher, 2009; Chang y Lin, 2011), que sirven para tareas de clasificación y regresión basados en datos de entrenamiento. La idea esencial es lograr que el algoritmo aprenda a partir de ejemplos (situaciones), de los cuales se conoce su salida correcta (clasificación); para ello los SVM separan los puntos (datos de entrenamiento)

en dos clases, de tal forma que se maximiza la distancia entre el hiperplano que separa las clases y los puntos más cercanos que pertenecen a estas.

Primero se debe entrenar a los SVM con situaciones cuya clasificación se conoce, esto genera un modelo de clasificación del problema (Fletcher, 2009). Luego, el modelo recibe una situación *x* cuya clasificación desconoce, el modelo entonces compara la situación *x* con situaciones cuya clasificación conoce y genera la clasificación para la situación *x*.

En las figuras 1 y 2 se muestra cómo se usaron los SVM para la clasificación emocional de las canciones.

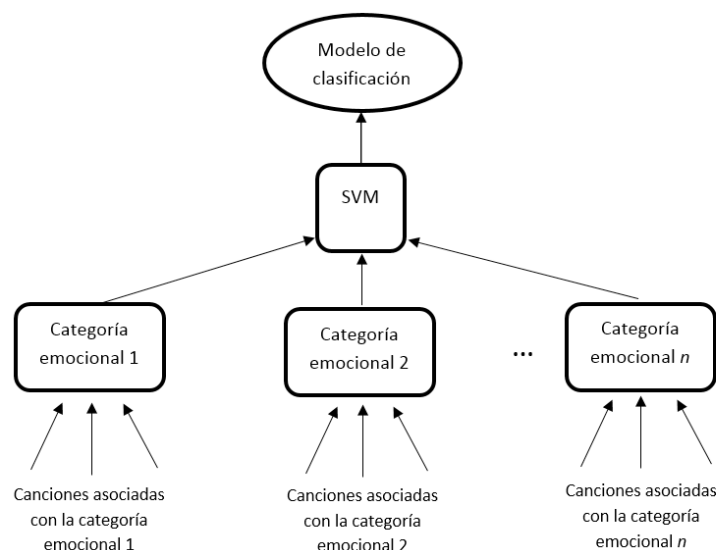


Figura 1. Modelo de clasificación para las canciones

Fuente: elaboración propia.

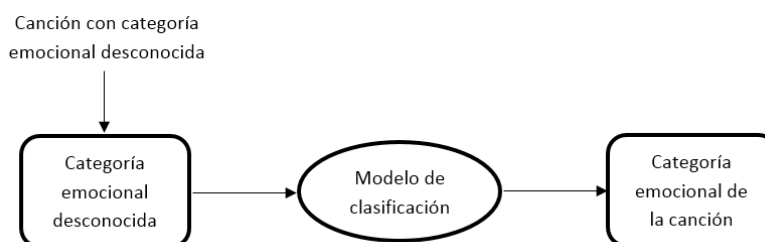


Figura 2. Aplicación del modelo de clasificación para una canción

Fuente: elaboración propia.

Dado que los SVM funcionan con parámetros numéricos, tanto en la entrada como en la salida, se debe convertir la letra en un formato válido de entrada para los SVM. Para ello se usó el TF-IDF (*term frequency – inverse document frequency*) (Ramos, 2003).

El TF-IDF es una medida numérica que expresa qué tan relevante es un término (palabra) para un documento con relación a un conjunto de documentos. Esta medida se utiliza a menudo como un factor de ponderación en la recuperación de información y en la minería de texto. Los términos con más alto TF-IDF son, por lo general, aquellos que mejor caracterizan la temática de un documento. Se calcula mediante la ecuación (1).

$$TFIDF(j) = \frac{t_j}{|d|} * \log \frac{N}{n_j} \quad (1)$$

Donde:

t_j : es el número de veces que aparece el término j en el documento d .

$|d|$: es el número de términos distintos que tiene el documento d .

N : es el número total de documentos.

n_j : es el número de documentos que contienen el término j .

Para el ejemplo en curso, considérese $j = feel$, $t_j = 1$, $|d| = 23$ (letra de la canción correspondiente al paso 3), $N = 10$ canciones y $n_j = 4$. Entonces $TFIDF(feel) = (1/23) * \log(10/4) = 0,01730174$.

De esta forma, a un término j se le asigna un peso en el documento d así: a) el TF-ID es alto cuando el término j ocurre muchas veces en un número pequeño de documentos, b) el TF-ID es bajo cuando el término j ocurre pocas veces en un documento y ocurre en muchos documentos y c) el TF-ID es muy bajo cuando el término j ocurre en todos los documentos.

El TF-IDF se suele normalizar, ya que el uso de esta frecuencia de los términos puede conducir a problemas como *keyword spamming*, es decir, la repetición de un término en un documento con el propósito de mejorar su clasificación en un

sistema de recuperación de información o la generación de un sesgo hacia los documentos largos, haciendo que se vean más importantes de lo que son debido a la alta frecuencia del término en el documento. Para normalizarlo, se obtiene el TF-IDF para cada término j del documento d , se suman los cuadrados de estos y se obtiene la raíz cuadrada de dicha sumatoria (ecuación (2)). El resultado pasa a dividir el TF-IDF del término j y ese será su TF-IDF *normalizado*. Para el ejemplo, la sumatoria fue 0,0808 y la raíz = 0,2844. Los resultados para cinco términos de la canción se muestran en la tabla 1.

$$TFIDF_{normalizado}(j) = \frac{TFIDF(j)}{\sqrt{\sum_{k=1}^{|d|} TFIDF(k)^2}} \quad (2)$$

Tabla 1. Resultados del TF-IDF para siete términos de la canción

Término(j)	t_j	n_j	TF-IDF	TF-IDF normalizado
bring	2	2	0,061	0,214
dream	1	2	0,031	0,107
feel	1	4	0,017	0,061
get	3	5	0,039	0,138
go	1	5	0,013	0,046

Fuente: elaboración propia.

Continuando con el paso 4, se consideró un conjunto SCAT de 18 categorías emocionales definidas en Hu, Downie y Ehmann (2009): 1) *aggression* (agresión), 2) *anger* (ira), 3) *angst* (angustia), 4) *broodiness* (desasosiego), 5) *calm* (calma), 6) *cheerfulness* (alegría), 7) *confidentiality*, (confianza), 8) *depression* (depresión), 9) *desire* (anhelo), 10) *dreaminess* (ensoñamiento), 11) *earnest* (seriedad), 12) *excitation* (excitación), 13) *grief* (dolor), 14) *happiness* (felicidad), 15) *pessimism* (pesimismo), 16) *romanticism* (romanticismo), 17) *sadness* (tristeza) y 18) *upbeatness* (optimismo).

Se seleccionaron 30 canciones (datos de entrenamiento) *para cada categoría*. La selección de

las canciones para cada una de las 18 subcategorías emocionales se hizo mediante la clasificación emocional proporcionada por sitios como *Stereo-mood.com*, *Musicoverly.com*, *Moodstream.gettyimages.com*, *Thesixtyone.com* y *Lastfm.es*. Estos sitios facilitaron, en gran medida, la clasificación, pues cuentan con grandes bases de datos que relacionan las canciones con sus respectivos estados emocionales. Sin embargo, como no todas las categorías emocionales utilizadas en estos sitios coincidían con las categorías utilizadas en el presente artículo, se estableció una correspondencia con las más similares. Además, para la selección de canciones de algunas de las categorías (como *Earnest*, *Confident* y *Brooding*) se tuvo que acudir a la opinión de usuarios en redes sociales y a búsquedas específicas mediante Google.

Por último, luego de obtener los valores TF-IDF para cada uno de los 23 términos de la canción, se ejecutó el modelo generado por los SVM, el cual a partir de los datos de entrenamiento clasificó la canción “Come on, get happy” en la categoría *happy*, lo cual concuerda con la clasificación usualmente dada a esta canción en los sitios mencionados.

METODOLOGÍA Y ALGORITMO PROPUESTO

El objetivo del algoritmo es *clasificar un álbum* en una categoría emocional basado en la clasificación emocional de cada una de sus canciones. Para ello, se consideró un conjunto *CAT* de seis categorías emocionales tomadas del modelo de Tellegen-Watson-Clark (Trohidis, K., Tsoumakas, G., Kalliris, G., & Vlahavas, 2011; Robertson, 2004): 1) *amazement-surprise* (asombro-sorpresa), 2) *happiness-satisfaction* (felicidad-satisfacción), 3. *relaxation-calmness* (relajación-calma), 4) *quietness-stillness* (tranquilidad-quietud), 5) *sadness-loneliness* (tristeza-soledad), 6) *anger-fear* (enojo-temor). Estas categorías agrupan al conjunto *SCAT* de 18 subcategorías en las cuales se clasifica una canción. Es decir, se consideran seis

categorías para clasificar un álbum y 18 subcategorías para clasificar una canción. Esto es razonable porque un álbum está compuesto de canciones y una canción tiende a estar orientada a una subcategoría emocional específica; por tanto, la categoría emocional de un álbum será una ponderación de las subcategorías de sus canciones. Por ejemplo, si un álbum tiene diez canciones, donde cinco son clasificadas en *depression*, tres en *grief* y dos en *broodiness*, se podría concluir que la tendencia predominante del álbum es de *sadness-loneliness*, que es una categoría que abarca a las tres subcategorías predominantes en sus canciones.

Con el fin de establecer el grado (peso) en que las subcategorías contribuyen con las categorías, se debe notar que algunas de las 18 subcategorías no se diferencian claramente de otras, por ejemplo, *cheerfulness*, *happinnes* y *upbeatness*. Además, una persona podría considerar, por ejemplo, que el dolor (subcategoría *grief*) conlleva principalmente a un estado de tristeza (categoría *sadness-loneliness*), y otra podría considerar que dicha subcategoría conlleva principalmente a un estado de ira (categoría *anger-fear*). Por tanto, existe un elemento de subjetividad que el analista musical puede incorporar en el modelo propuesto.

Para afrontar este aspecto, se definió una matriz de pesos emocionales W , la cual indica cuán importante (peso) es una subcategoría para una categoría. Esta matriz es *parametrizable por parte del analista musical*, ya que para alguien muy melancólico, la tristeza puede en cierto grado dar felicidad; sin embargo, existe una restricción:

$$\sum_{i=1}^l w(j, i) = 1, 1 \leq j \leq k, \text{ donde } k \text{ es el número de categorías (6) y } l \text{ es el número de subcategorías (18).}$$
Es decir, la sumatoria de los pesos de las subcategorías que aportan a una categoría debe ser igual a 1. Se estableció además que una categoría está compuesta por un subconjunto de las 18 subcategorías. Por ejemplo, en la tabla 2 la categoría *amazement-surprised* (1) está compuesta por 70 % de la subcategoría *excitation* (12) y 30 % de la subcategoría *happiness* (14).

Tabla 2. Matriz de pesos emocionales W

Cat.	Subcategoría																		Σ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
1												0,7		0,3					1
2						0,3								0,5		0,1		0,1	1
3					0,4	0,1	0,1		0,1	0,1	0,1			0,1					1
4				0,1	0,4		0,2				0,3								1
5			0,2					0,2					0,2		0,2		0,2		1
6	0,3	0,3						0,3									0,1		1

Fuente: elaboración propia.

Considerando lo anterior se propone el algoritmo *emoAlbumClass1*. El algoritmo *emoAlbumClass1* calcula el estado emocional de un álbum basado en la subcategoría emocional de cada una de sus n canciones. Es decir, el algoritmo calcula el peso que obtiene cada álbum por categoría y lo clasifica en la categoría en donde obtenga mayor peso.

Algoritmo *emoAlbumClass1*

Input:

$CAT[1..k]$ //Arreglo de nombres de las k categorías para clasificar un álbum

$SCAT[1..l]$ //Arreglo de nombres de las l subcategorías para clasificar una canción

$CA[1..n]$ /*Arreglo de códigos (correspondientes a las posiciones de SCAT) de las subcategorías para las n canciones de un álbum */

$W[1..k, 1..l], W[i][j]$ in $[0,1], i, j: 1 \leq i \leq k, 1 \leq j \leq l$
//Matriz de pesos

Output:

$CAT[p]$ //Categoría a la que pertenece el álbum,
 $1 \leq p \leq k$

Begin

1. $R[1..k], R[i] = 0, i: 1 \leq i \leq k$ // $R[i]$ corresponde a $CAT[i]$.

2. **For** $i = 1$ **to** k **Do**

3. **For** $j = 1$ **to** n **Do**

4. $R[i] = R[i] + W[i, CA[j]]$ /*Sumatoria de los pesos aportados por la subcategoría $CA[j]$ de cada canción a la categoría $R[i]$ */

5. **End For**

6. **End For**

7. $max =$ elemento p en R con mayor valor $R[p]$,
 $1 \leq p \leq k$

8. **Return** $CAT[max]$ //Retorna la categoría que obtuvo el máximo valor en R

End

A continuación se propone el algoritmo *emoAlbumClass2*. El algoritmo *emoAlbumClass2*, al igual que el algoritmo *emoAlbumClass1*, calcula el estado emocional de un álbum basado en el estado emocional de sus canciones, con la diferencia que el algoritmo *emoAlbumClass2* considera adicionalmente la subcategoría secundaria emocional de cada canción generada por los SVM.

Algoritmo *emoAlbumClass2*

//Considera la subcategoría principal y secundaria de cada canción

Input:

$CAT[1..k]$ //Arreglo de nombres de las k categorías para clasificar un álbum

$SCAT[1..l]$ //Arreglo de nombres de las l subcategorías para clasificar una canción

$CA1[1..n]$ /*Arreglo de códigos (correspondientes a las posiciones de SCAT) de las subcategorías principales para las n canciones de un álbum */

$CA2[1..n]$ /*Arreglo con los códigos (correspondientes a las posiciones de SCAT) de las subcategorías secundarias para las n canciones de un álbum */
 $W[1..k, 1..l], W[i][j]$ in $[0, 1], i, j: 1 \leq i \leq k, 1 \leq j \leq l$
//Matriz de pesos

Output:

$CAT[p]$ //Categoría a la que pertenece el álbum,
 $1 \leq p \leq k$
Begin
 1. $R[1..k]$, $R[i] = 0$, $i: 1 \leq i \leq k$ // $R[i]$ corresponde a $CAT[i]$.
 2. **For** $i = 1$ **to** k **Do**
 3. **For** $j = 1$ **to** n **Do**
 4. $R[i] = R[i] + W[i, CA1[j]] + W[i, CA2[j]]$ /* Sumatoria de los pesos aportados por las subcategorías $CA1[j]$ y $CA2[j]$ de cada canción a la categoría $R[i]$ */
 5. **End For**
 6. **End For**
 7. max = elemento p en R con mayor valor $R[p]$, $1 \leq p \leq k$
 8. **Return** $CAT[max]$ //Retorna la categoría que obtuvo el máximo valor en R
End

Ejemplo: considérense las canciones del álbum *Slippery When Wet*, de Bon Jovi, cada una con su subcategoría principal y secundaria como se muestra en la tabla 3.

Nótese que las subcategorías principal y secundaria de cada canción son similares. Esto sugiere que cada canción presenta una tendencia hacia una emoción en particular.

a) Aplicación del algoritmo *emoAlbumClass1*. Sean: $CAT = [Amazement-surprise, Happiness-satisfaction, Relaxation-calmness, Quietness-stillness, Sadness-loneliness, Anger-fear]$.

$SCAT = [Aggression, Anger, Angst, Broodiness, Calm, Cheerfulness, Confidentiality, Depression, Desire, Dreaminess, Earnest, Excitation, Grief, Happiness, Pessimism, Romanticism, Sadness, Upbeatness]$.

$CA = [11, 6, 8, 13, 11, 6, 15, 15, 16, 15]$.

Y sea W la matriz de pesos de la tabla 4. El resultado para la categoría ganadora es:

$R[5] = W(5, CA[1]) + W(5, CA[2]) + \dots + W(5, CA[10]) = 0 + 0 + 0,2 + 0,2 + 0 + 0 + 0,2 + 0,2 + 0 + 0,2 = 1$.

Los resultados se muestran en la tabla 4, en gris se muestra la categoría ganadora *sadness-loneliness* (5).

b) Aplicación del algoritmo *emoAlbumClass2*.

Sean CAT , $SCAT$ y W como en el caso a), $CA1 = [11, 6, 8, 13, 11, 6, 15, 15, 16, 15]$ y $CA2 = [8, 14, 11, 11, 3, 9, 3, 3, 10, 8]$.

El resultado para la categoría ganadora es:

$R[5] = W(5, CA1[1]) + W(5, CA1[2]) + \dots + W(5, CA1[10]) + W(5, CA2[1]) + W(5, CA2[2]) + \dots + W(5, CA2[10]) = 2$.

Tabla 3. Subcategoría principal y secundaria de cada canción del álbum *Slippery When Wet*.

Canción (autores)	Subcategoría principal	Subcategoría secundaria
"I'd Die For You" (J. Bon Jovi, R. Sambora)	<i>Earnest</i> (11)	<i>Depression</i> (8)
"Let It Rock" (J. Bon Jovi, R. Sambora)	<i>Cheerfulness</i> (6)	<i>Happiness</i> (14)
"Livin' On A Prayer" (J. Bon Jovi, R. Sambora, D. Child)	<i>Depression</i> (8)	<i>Earnest</i> (11)
"Never Say Goodbye" (J. Bon Jovi, R. Sambora)	<i>Grief</i> (13)	<i>Earnest</i> (11)
"Raise Your Hands" (J. Bon Jovi, R. Sambora)	<i>Earnest</i> (11)	<i>Angst</i> (3)
"Social Disease" (J. Bon Jovi, R. Sambora)	<i>Cheerfulness</i> (6)	<i>Desire</i> (9)
"Wanted Dead Or Alive" (J. Bon Jovi, R. Sambora)	<i>Pessimism</i> (15)	<i>Angst</i> (3)
"Wild In The Streets" (J. Bon Jovi)	<i>Pessimism</i> (15)	<i>Angst</i> (3)
"Without Love" (J. Bon Jovi, R. Sambora, D. Child)	<i>Romanticism</i> (16)	<i>Dreaminess</i> (10)
"You Give Love A Bad Name" (J. Bon Jovi, R. Sambora, D. Child)	<i>Pessimism</i> (15)	<i>Depression</i> (8)

Fuente: elaboración propia.

Los resultados se muestran en la tabla 5, en gris se muestra la categoría ganadora, de nuevo *sadness-loneliness* (5).

Nótese que los números subrayados representan el aporte de la subcategoría principal de cada canción a la categoría i , por ejemplo, $W[i, CA1[j]]$; mientras que los números sin subrayar representan el aporte de la subcategoría secundaria j de cada canción a la categoría i , por ejemplo, $W[i, CA2[j]]$.

EXPERIMENTOS Y RESULTADOS

En una primera fase de la verificación experimental, se trabajó con una base de datos de 540 canciones para generar el modelo de clasificación de estas en las 18 subcategorías, es decir, 30 canciones por cada subcategoría. Luego se tomó un conjunto de prueba de otras 15 canciones por cada subcategoría (270 canciones) y se logró un grado de precisión en la clasificación del 65 %. Para el procesamiento de la letra de cada canción (pasos en “Fundamentos teóricos”) se usó *Rapid*

Miner, un *software* especializado para minería de datos. Los algoritmos *emoAlbumClass1* y *emoAlbumClass2* se implementaron en *Java*.

A continuación se seleccionaron seis artistas cada uno con cuatro álbumes (24 álbumes en total). Estos últimos están conformados por canciones entre las 270 ya clasificadas. Con el fin de obtener la clasificación emocional de cada álbum, se aplicaron los algoritmos *emoAlbumClass1* y *emoAlbumClass2*.

Además, se consideró una tercera estrategia para obtener la clasificación emocional de un álbum: se tomó la letra de cada una de sus canciones y se unieron como si formasen la letra de una sola *canción*. Se generó entonces un segundo modelo de clasificación de las canciones en las seis categorías del conjunto CAT. Se trabajó con la misma base de datos de 540 canciones para la fase de entrenamiento y 270 para la fase de prueba. Se logró un porcentaje de precisión en la clasificación de las canciones del 82 %, esto es razonable porque ahora el número de categorías para

Tabla 4. Ejemplo del algoritmo *emoAlbumClass1*

Cat.	Canción										Σ
	1	2	3	4	5	6	7	8	9	10	
1											0
2		0,3				0,3			0,1		0,7
3	0,1	0,1			0,1	0,1					0,4
4	0,3				0,3						0,6
5			0,2	0,2			0,2	0,2		0,2	1
6			0,3								0,3

Fuente: elaboración propia.

Tabla 5. Ejemplo del algoritmo *emoAlbumClass2*

Cat.	Canción										Σ
	1	2	3	4	5	6	7	8	9	10	
1		<u>0</u> +0,3									0,3
2		<u>0,3</u> +0,5	<u>0</u> +0,1	<u>0</u> +0,1		<u>0,3</u> +0			<u>0,1</u> +0		1,4
3	<u>0,1</u> +0	<u>0,1</u> +0,1	<u>0</u> +0,3	<u>0</u> +0,3	<u>0,1</u> +0	<u>0,1</u> +0,1			<u>0</u> +0,1		1,2
4	0,3				0,3+0						0,6
5	0+0,2		0,2+0	0,2+0	0,2		0,2+0,2	0,2+0,2		0,2+0,2	2
6	<u>0</u> +0,3		<u>0,3</u> +0								0,6

Fuente: elaboración propia.

clasificar las canciones era solo seis. Se usó entonces este segundo modelo de clasificación para obtener la clasificación emocional del álbum (el parámetro de entrada para el modelo de clasificación fue la unión de las letras de todas las canciones del álbum).

Los experimentos se ejecutaron en un computador portátil con 8 GB de memoria RAM y procesador Core i7. La clasificación emocional de un álbum se obtuvo en menos de un minuto, tiempo

razonable para hacer análisis en línea. Los resultados se muestran en la tabla 6.

En la tabla 6 se observa que la categoría generada por los algoritmos *emoAlbumClass1* y *emoAlbumClass2* tiende a ser igual en 22 de los 24 álbumes (92 %), esto sugiere que estos álbumes tienen un estado emocional predominante, lo que se refleja en la mayoría de sus canciones. Sin embargo, también se analizan dos álbumes donde la categoría generada por estos dos algoritmos es

Tabla 6. Resultados de los experimentos

Artista	Álbum y año	Categoría emocional aplicando el algoritmo <i>emoAlbumClass1</i>	Categoría emocional aplicando el algoritmo <i>emoAlbumClass2</i>	Categoría emocional aplicando la unión de las letras de todas las canciones del álbum	Observaciones
Bon Jovi	<i>Slippery when Wet</i> [1986]	Happiness-satisfaction		Happiness-satisfaction	Los álbumes del artista tienden a ser emocionalmente alegres.
	<i>New Jersey</i> [1988]	Happiness-satisfaction		Quietness-stillness	
	<i>These Days</i> [1995]	Happiness-satisfaction		Happiness-satisfaction	
	<i>Bounce</i> [2002]	Happiness-satisfaction		Quietness-stillness	
Justin Bieber	<i>My World</i> [2009]	Relaxation-calmness		Quietness-stillness	Los álbumes del artista tienden a ser emocionalmente calmados y alegres en menor grado.
	<i>My World 2.0</i> [2010]	Quietness-stillness		Quietness-stillness	
	<i>Under The Mistletoe</i> [2011]	Quietness-stillness		Quietness-stillness	
	<i>Believe</i> [2012]	Happiness-satisfaction	Quietness-stillness	Happiness-satisfaction	
Eminem	<i>Infinite</i> [1996]	Anger-fear		Anger-fear	Los álbumes del artista tienden a ser emocionalmente agresivos.
	<i>The Marshall Mathers LP</i> [2000]	Anger-fear		Anger-fear	
	<i>The Eminem Show</i> [2002]	Anger-fear		Anger-fear	
	<i>Bad Meets Evil–Hell The Sequel</i> [2011]	Anger-fear		Anger-fear	
	<i>Off the Wall</i> [1979]	Amazement-surprise		Amazement-surprise	
Michael Jackson	<i>Thriller</i> [1982]	Relaxation-calmness		Relaxation-calmness	Los álbumes del artista <i>no</i> tienden a un estado emocional específico. Se observan estados emocionales que incluyen asombro, alegría y calma.
	<i>Bad</i> [1987]	Happiness-satisfaction		Happiness-satisfaction	
	<i>Dangerous</i> [1991]	Amazement-surprise	Relaxation-calmness	Relaxation-calmness	
	<i>Bleach</i> [1989]	Relaxation-calmness		Relaxation-calmness	
	<i>Nevermind</i> [1991]	Quietness-stillness		Quietness-stillness	
Nirvana	<i>In Utero</i> [1993]	Sadness-loneliness		Sadness-loneliness	Los álbumes del artista tienden a ser emocionalmente melancólicos.
	<i>From the Muddy Banks of Wishkah</i> [1996]	Sadness-loneliness		Quietness-stillness	
	<i>A Night at the Opera</i> [1975]	Amazement-surprise		Amazement-surprise	
Queen	<i>A Day at the Races</i> [1976]	Happiness-satisfaction		Quietness-stillness	Los álbumes del artista tienden a ser emocionalmente alegres.
	<i>News of the World</i> [1977]	Happiness-satisfaction		Happiness-satisfaction	
	<i>Innuendo</i> [1991]	Happiness-satisfaction		Happiness-satisfaction	

Fuente: elaboración propia.

diferente (celdas sombreadas en la tabla 6). Esto sugiere que la categoría emocional del álbum no es tan predominante y donde la subcategoría secundaria de sus canciones fue decisiva para la clasificación. También, en la mayoría de los casos, un artista tiende a mantener a través del tiempo la misma categoría emocional en sus álbumes, es decir, hay una *consistencia emocional* a través del tiempo en su producción discográfica.

Por otro lado, al comparar los resultados de estos dos algoritmos con la tercera alternativa (unión de las letras de todas las canciones del álbum): a) se obtuvo el mismo resultado en 17 de los 24 álbumes (70 %), b) en los dos álbumes donde no coincidió la categoría emocional generada por los algoritmos *emoAlbumClass1* y *emoAlbumClass2*, una de estas dos categorías sí coincidió con la categoría generada por la tercera alternativa, y c) se obtuvo diferente resultado en 5 de los 24 álbumes.

El hecho de registrar en un 70 % de los álbumes el mismo resultado, mediante tres estrategias, indica un alto grado de consistencia entre ellas y sugiere en la mayoría de los álbumes una tendencia emocional dominante. Para los otros álbumes (30 %) la diferente clasificación lograda se puede explicar por álbumes que no tienen una tendencia emocional dominante y por sesgos e imprecisiones inherentes a los modelos generados por los SVM.

En un segundo experimento, se seleccionaron diez artistas: AC/DC, Blink-182, Bon Jovi, Bob Marley, Britney Spears, Daft Punk, Eminem, Green Day, Metallica y Queen. De cada artista se seleccionaron ocho álbumes y de cada álbum cuatro canciones, para lo cual, se analizaron 80 álbumes y 320 canciones. Para cada artista se halló el número máximo de álbumes que pertenecen a una misma categoría emocional (se aplicó el algoritmo *emoAlbumClass1*). A este valor, expresado en porcentaje (con respecto a sus ocho álbumes), se le denominará *consistencia emocional*. La tabla 7 muestra la consistencia emocional para los artistas analizados.

Tabla 7. Consistencia emocional de los álbumes de los artistas

Cantidad de artistas	Número de álbumes en una misma categoría emocional	Consistencia emocional
2	7	87,5 %
2	6	75 %
4	5	62,5 %
1	4	50 %
1	3	37,5 %

Fuente: elaboración propia.

Los resultados indican que ocho de los diez artistas tuvieron una consistencia emocional superior al 50 %. Esto sugiere que un artista tiende a mantener en sus álbumes la misma categoría emocional a través del tiempo.

Por último, en un tercer experimento, se seleccionaron diez géneros musicales; de cada género, cuatro artistas representativos; de cada artista, tres álbumes, y de cada álbum, cuatro canciones; así, se analizaron 120 álbumes y 480 canciones. En la tabla 8 se muestran los géneros seleccionados y sus artistas respectivos.

Tabla 8. Artistas por género

Género	Artistas
Country	Blake Shelton, George Strait, Johnny Cash, Tim McGraw
Electrónica	Armin Van Buuren, Daft Punk, David Guetta, Faithless
Gospel	Deitrick Haddon, Lecrae, Mary Mary, Tamela Mann
Hip hop	2Pac, Eminem, Lil Wayne, Snoop Dogg
Metal	Black Sabbath, Iron Maiden, Judas Priest, Metallica
Pop	Britney Spears, Bruno Mars, Katy Perry, Lady Gaga
Punk	Blink-182, Green Day, The Clash, The Ramones
Reggae	Bob Marley, Jimmy Cliff, Peter Tosh, UB 40
Rock	AC/DC, Bon Jovi, Queen, Scorpions
Soul	Aretha Franklyn, Gladys Knight, Marvin Gaye, Otis Redding

Fuente: elaboración propia.

La tabla 9 muestra el número de álbumes de cada género clasificados por categoría emocional (se aplicó el algoritmo *emoAlbumClass1*).

Aunque se requieren experimentos más exhaustivos, los resultados evidenciaron algunos aspectos interesantes. Por ejemplo, en géneros como el *punk* y el *metal*, usualmente considerados agresivos y depresivos, solo la mitad de sus álbumes se clasificaron en las categorías emocionales *anger-fear* y *sadness-loneliness*. Por su parte, el género *rap*, cuyas letras usualmente conllevan crítica social y palabras fuertes, tuvo más de la mitad de sus álbumes en estas dos categorías. Por otro lado, sorprende que en el género electrónica haya tres álbumes en la categoría *anger-fear* ya que este género, por lo general, está orientado a letras alegres y “animadas”. En cuanto a los géneros *gospel* y *reggae*, más de la mitad de sus álbumes se ubicaron entre las categorías *happiness-satisfaction*, *relaxation-calmness* y *quietness-stillness*; esto concuerda con sus letras alegres, optimistas y que suelen incluir mensajes de paz y amor. En el resto de los géneros no se observó una tendencia en cuanto a la categoría emocional de sus álbumes.

CONCLUSIONES

En este artículo se propusieron dos algoritmos para clasificar un álbum en una categoría emocional. En ambos, la clasificación tuvo en cuenta la clasificación emocional de cada una de las canciones del álbum, para ello se usaron los SVM. La diferencia entre los dos algoritmos es que uno considera solo la categoría emocional principal asociada con cada canción; mientras que el otro, su categoría emocional secundaria. También se propuso una tercera estrategia para determinar la categoría emocional de un álbum, la cual considera la unión de las letras de todas las canciones del álbum.

Para la clasificación de los álbumes se tuvieron en cuenta seis categorías y para las canciones 18 subcategorías. Así, las categorías emocionales de los álbumes abarcan las subcategorías de las canciones. Esto es razonable porque, por lo general, una canción se enfoca en una emoción particular, mientras que un álbum suele ser más heterogéneo desde el punto de vista emocional. Por consiguiente, la clasificación de un álbum es una ponderación de las emociones de cada una de sus canciones. Una de las ventajas de la propuesta es

Tabla 9. Álbumes por género

Género	Categoría emocional					
	Amazement-surprise	Happiness-satisfaction	Relaxation-calmness	Quietness-stillness	Sadness-loneliness	Anger-fear
Country	0	5	4	2	1	0
Punk	4	2	0	0	2	4
Pop	3	3	2	1	2	1
Gospel	0	6	3	3	0	0
Metal	4	2	1	0	2	3
Soul	3	4	2	2	1	0
Rap	3	2	0	0	2	5
Electrónica	5	3	1	0	0	3
Rock	2	4	2	2	1	1
Reggae	2	6	2	2	0	0
Total	26	37	17	12	11	17

Fuente: elaboración propia.

que el analista puede especificar el grado (peso) con que cada subcategoría contribuye a una categoría, lo que permite introducir un elemento de subjetividad en el modelo. Para el 70 % de los álbumes analizados se obtuvo la misma categoría emocional mediante las tres estrategias, esto indica un alto grado de consistencia entre las propuestas.

Por otro lado, se analizó la consistencia emocional de los álbumes de algunos artistas; esto es, si un artista a través del tiempo mantiene la misma categoría emocional en sus álbumes. Los resultados mostraron que el 80 % de los artistas analizados mantienen una consistencia emocional superior al 50 %. También se analizó la correlación entre las categorías emocionales y algunos géneros musicales, aunque se requieren experimentos más exhaustivos, los resultados mostraron que ciertos géneros (como el *metal* y el *punk*), que tradicionalmente son asociados con ciertas emociones (ira, depresión) no necesariamente exhiben tal tendencia.

Una de las desventajas de la propuesta es que solo funciona para canciones que tengan letra, así no es posible clasificar álbumes que son parcial o totalmente instrumentales. Para abordar este aspecto se deben considerar aspectos de audio, los cuales se espera incorporar en un trabajo futuro. Esto permitiría identificar, por ejemplo, la categoría emocional de un segmento de una sinfonía y establecer su correspondencia con los *tempos* (como *adagio*, *andante* y *moderato*, entre otros). Otra desventaja es que no se consideran figuras retóricas que afectan el significado y sentido literario como la ironía, el eufemismo y la sinestesia, entre muchas otras y que son frecuentes en las letras de las canciones. Tampoco se consideran los tiempos verbales (por ejemplo, “*I was very sad but now I am very happy*”) ni negaciones (“*I won’t be sad for you*”). También, existen letras cuyo significado es ambiguo, poético o que exige un análisis detallado (por ejemplo, “*Said I loved you but I lied ‘Cause this is more than love I feel inside, Said I loved you but I was wrong ‘Cause love could never ever feel so strong*”). Estas y otras situaciones ofrecen

posibilidades de trabajos futuros con el fin de mejorar los algoritmos propuestos o proponer otros.

Otro trabajo futuro es analizar si los aspectos de audio mejoran significativamente la clasificación de las canciones que incluyen letras. También, se debe profundizar en algunos géneros y considerar sus subgéneros (para el caso del *metal* tener en cuenta subgéneros como *black*, *gothic* y *trash*, entre muchos otros) con relación a las categorías emocionales. Esto podría explicar por qué en algunos géneros se presenta tanta variedad de categorías emocionales. Se podría considerar el uso de sinónimos, antónimos, términos polisémicos, relaciones jerárquicas entre términos, entre otros. Estos aspectos sugieren el uso de ontologías que podrían enriquecer la propuesta y mejorar los resultados de la clasificación. Por último, se podrían aprovechar otras características del álbum para tratar de deducir u obtener al menos un indicio sobre su categoría emocional, por ejemplo, la portada de algunos álbumes es en algunas ocasiones sugerente (portadas de los álbumes de bandas de *metal* como Iron Maiden con su monstruo Eddie o las portadas de álbumes de *trance* que suelen incluir paisajes paradisíacos).

FINANCIAMIENTO

Universidad Nacional de Colombia sede Medellín.

REFERENCIAS

- Agichtein, E. & Gravano, L. (2000). Snowball: Extracting Relations from Large Plain-Text Collections. En: *Fifth ACM Conference on Digital Libraries* (pp. 85–94). Nueva York: ACM.
- Chang, C. & Lin, C. (2011). LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology* 2(3), 1-27.
- Cherman, E.; Monard, M. y Metz, J. (2011). Multi-Label Problem Transformation Methods: A Case Study. *CLEI Electronic Journal* 14(1), 1-10.
- Diesner, J. (2006). Part of Speech Tagging for English Text Data. *Machine Learning Project Reports*, 1-8.

- Recuperado de: <http://www.cs.cmu.edu/~epxing/Class/10701-06f/project-reports/diesner.pdf>
- Feldman, R. & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge: Cambridge University Press.
- Ferragina, P. & Scaiella, U. (2011). First Steps Beyond the Bag-of-Words Representation of Short Texts. *Italian Information Retrieval Workshop* (pp. 1-4). Milán: Universidad de Milán.
- Fletcher, T. (2009). Support Vector Machines Explained Acknowledgments. *The Scientific World Journal*, 1-19. Recuperado de: <http://www.tristanfletcher.co.uk/SVM%20Explained.pdf>
- Hu, X.; Downie, J. & Ehmann, A. (2009). Lyric Text Mining in Music Mood Classification. En: *10th International Society for Music Information Retrieval Conference* (pp. 411-416). Utrecht: Ismir.
- Jivani, A. (2011). A Comparative Study of Stemming Algorithms. *International Journal of Computer Technology and Applications* 2(6), 1930-1938.
- Li, T. & Ogihara, M. (2003). Detecting Emotion in Music. En: *4th international conference on music information retrieval* (pp. 239-240). Baltimore: Ismir.
- Liu, T.; Moore, A.; Gray, A. & Yang, K. (2004). An Investigation of Practical Approximate Nearest Neighbor Algorithms. En: *Proceedings of Advances in Neural Information Processing Systems*, 17. Recuperado de: <http://papers.nips.cc/paper/2666-an-investigation-of-practical-approximate-nearest-neighbor-algorithms.pdf>
- Noraida, H. & Noor, I. (2012). Porter Stemming Algorithm for Semantic Checking. En: *International Conference on Computing and Information Technology* (pp. 253-258). Amman: IEEE.
- Panagakos, Y. & Kotropoulos, C. (2011). Automatic Music Mood Classification Via Low-Rank Representation. En: *19th European Signal Processing Conference* (pp. 689-693). Cataluña: Universidad Politécnica de Cataluña.
- Ramos, J. (2003). Using TF-IDF to Determine Word Relevance in Document Queries. En: *First Instructional Conference on Machine Learning* (pp. 5-8). Piscataway: Rutgers University.
- Read, J.; Pfahringer, B.; Holme, G. & Frank, E. (2011). Classifier chains for multi-label classification. *Machine Learning* 85(3), 333-359.
- Roberts, A. (2002). *Automatic Acquisition of Word Classification Using Distribution Analysis of Content Words with Respect to Function Words*. Leeds: University of Leeds.
- Robertson, S. (2004). Understanding Inverse Document Frequency: On Theoretical Arguments for IDF. *Journal of Documentation* 60(5), 503-520.
- Sharma, D. (2012). Stemming Algorithms: A Comparative Study and their Analysis. *International Journal of Applied Information Systems* 4(3), 7-12.
- Trohidis, K., Tsoumakas, G., Kalliris, G., & Vlahavas, I. (2011). Multi-Label Classification of Music by Emotion. *Journal on Audio, Speech, and Music Processing* 4(1), 325-330.
- Tsoumakas, G.; Katakis, I. & Vlahavas, I. (2011). Random k-Labelsets for Multilabel Classification. *Knowledge and Data Engineering* 23(7), 1079-1089.
- Zhang, L.; Zhou, W.; Chang, P.; Liu, J.; Yan, Z.; Wang, T. & Li, F. (2012). Kernel Sparse Representation-Based Classifier. *IEEE Transactions on Signal Processing* 60(4), 1684-1695.

