

Trabajo Fin de Grado

Estudio y elaboración de un sistema de
reconocimiento de emociones musicales mediante
minería de datos

Autor/es

Guillermo Ramos Sanz

Director/es

José Ramón Beltrán Blázquez, Piedad Garrido Picazo

Escuela Universitaria Politécnica de Teruel

2017

Sistema de reconocimiento de emociones musicales mediante minería de datos

Resumen

En el presente documento se presentará el Trabajo Final de Grado llevado a cabo por el alumno Guillermo Ramos Sanz en la Escuela Universitaria Politécnica de Teruel (EUPT). Dicho TFG muestra por un lado la investigación sobre la posibilidad de reconocer de forma automática emociones intrínsecas en una canción, y posterior realización de un sistema sencillo de apoyo a la toma de decisiones.

La música forma una parte fundamental en la vida de las personas. Uno de los mayores alicientes por los que la música es querida por los oyentes es por la capacidad de cambiar el estado anímico de las personas. Es innegable que existe relación entre música y emociones, por eso en este trabajo se va a profundizar en esta cuestión.

Existen varias formas de catalogar una emoción, usualmente divididas en categóricas o dimensionales. El primer paso es identificar las emociones y encontrar una catalogación que se considere adecuada para el TFG. A continuación, se investiga hasta qué punto se ha avanzado en las técnicas de reconocimiento automático de características en canciones, y cuáles son los parámetros primordiales que se están utilizando como referencia en otros proyectos similares.

Una vez entendidos esos puntos principales, es posible plantear un caso práctico, analizar el estado actual de progreso y proponer un sistema de apoyo que suponga un avance tecnológico. De esa forma se propondrá la realización de un sistema que asocie diferentes parámetros musicales no realizados hasta el momento, y que pueda tener una visión práctica.

Se diseña la parte práctica del proyecto, para posteriormente implementarlo. Se trata de un sistema libre que accede a varios servicios en línea para recopilar información de diferentes tipos y la unifica en el mismo repositorio. Esta información sirve de base tanto para la realización de estudios en profundidad como para nuevas asociaciones y relaciones.

Por último, para dotar de un carácter concluyente al proyecto se realizará un proceso sencillo de minado de datos, teniendo en cuenta que el sistema ha sido diseñado para poder trabajar con algoritmos como regresión lineal o de predicción. Este proceso nos permitirá sacar nuestras propias conclusiones y determinar si el enfoque planteado tiene sentido y es viable para continuar con la investigación llevada a cabo.

Palabras clave: Música, emociones, minería de datos, aprendizaje automático, sistema de ayuda a la toma de decisiones.

Contenido

Agradecimientos	6
1- Introducción	7
2- Estado del Arte	8
2.1- Emociones evocadas por la música: Caracterización, clasificación y medida	8
2.1.1- El modelo de emociones actual	8
2.1.2- Investigación previa en emociones musicales	8
2.1.3- Estudios previos interesantes:	9
2.1.4- Valoraciones.....	13
2.2- Reconocimiento automático de la emoción musical	14
2.2.1- Conceptualización de la emoción	15
2.2.2- Rasgos musicales.....	15
2.2.3- Reconocimiento categórico	16
2.2.4- Colección de datos	16
2.2.5- Preprocesamiento de datos	16
2.2.6- Anotación subjetiva	16
2.2.7- Entrenamiento del modelo	17
2.2.8- Reconocimiento dimensional.....	17
2.2.9- Detección de la variación en la emoción musical	17
2.2.10- Desafíos.....	18
2.2.11- Hueco semántico entre el rasgo musical de bajo nivel y la emoción humana de alto nivel	18
2.2.12- Explotación del timbre vocal.....	19
2.2.13- Conectar los sistemas MER dimensionales y categóricos.....	19
2.2.14- Considerar los factores situacionales.....	19
2.3- Maneras de reconocer emociones en música	19
2.3.1- Investigación psicológica.....	19
2.3.2- Entorno de trabajo para reconocimiento de emociones	21
2.3.3- Análisis del contenido del audio	22
2.3.4- Combinando múltiples análisis distintos	23
2.4- Software y tecnologías trabajando hoy en día en este ámbito	23
3- Plan de trabajo	25
4- Análisis.....	26
4.1- Requisitos.....	26
4.2- Sistema Operativo.....	26
4.3- Software de virtualización	27

4.4- Lenguaje de programación	27
4.5- Sistema gestor de base de datos	28
4.6- Fases.....	28
4.6.1- Primera fase: Acceso a datos multimedia.....	28
4.6.2- Segunda fase: Almacenamiento y validación de los datos	29
4.6.3- Tercera fase: Análisis y minería de datos.....	29
5- Diseño.....	30
5.1- Interfaces	30
5.2- Despliegue.....	30
5.3- Diagrama de actividad	30
5.4- Servicios externos	31
5.4.1- Spotify	32
5.4.2- Last.fm.....	33
5.4.3- Indico.io.....	33
5.5- Base de datos	34
6- Implementación	35
6.1- Gestión temporal	35
6.2- Decisiones de implementación.....	35
6.3- La ejecución determinante	38
7- Minería de datos	40
8- Licencia software y documental.....	45
9- Conclusiones y Trabajo Futuro	46
10- Referencias Bibliográficas	48

ÍNDICE DE ILUSTRACIONES

ILUSTRACIÓN 1: MUESTRA DE VARIACIÓN EN LA DETECCIÓN DE EMOCIONES. COMBINANDO LOS VALORES EN VALENCIA Y EXCITACIÓN PODEMOS FORMAR LA CURVA AFECTIVA (DERECHA), LA CUAL REPRESENTA LOS CAMBIOS DINÁMICOS DEL CONTENIDO AFECTIVO DE UNA PIEZA MUSICAL	17
ILUSTRACIÓN 2: ANOTACIONES EMOCIONALES PARA 4 CANCIONES: (A) SMELLS LIKE TEEN SPIRIT DE NIRVANA, (B) A WHOLE NEW WORLD DE PEABO BRYSON Y REGINA BELLE, (C) THE ROSE DE JANIS JOPLIN, (D) TELL LAURA I LOVE HER DE RITCHIE VALENS. NÓTESE COMO LA PERCEPCIÓN ES SUBJETIVA.....	18
ILUSTRACIÓN 3: CLASIFICACIÓN SEGÚN EL MODELO VALENCE-AROUSAL	20
ILUSTRACIÓN 4: MOODSWINGS, JUEGO CON PROPÓSITO DE AYUDAR AL ENTENDIMIENTO DE LAS EMOCIONES EN LA MÚSICA.....	21
ILUSTRACIÓN 5: DIAGRAMA DE ACTIVIDAD DE SME	31
ILUSTRACIÓN 6:ANÁLISIS PRELIMINAR DE LOS PARÁMETROS A ESTUDIAR	41
ILUSTRACIÓN 7: CLUSTERIZACIÓN QUE AGRUPA BAILABILIDAD CON MEDIDA RÍTMICA (TIMESIG, ABREVIACIÓN DE “TIME SIGNATURE”)	42
ILUSTRACIÓN 8:CLUSTERIZACIÓN ENTRE ALEGRÍA Y TEMPO.....	42
ILUSTRACIÓN 9: COMPARACIÓN ENTRE LOS RESULTADOS DE ENFADO, MIEDO Y SORPRESA RESPECTIVAMENTE	43
ILUSTRACIÓN 10: VALORES DE TRISTEZA SEGÚN EL MODO MUSICAL.....	43

ÍNDICE DE TABLAS

TABLA 1: FACTORES EXTRAÍDOS, LOS ADJETIVOS MÁS DISCRIMINATORIOS Y LAS ALFAS DE CRONBACH PARA LAS TRES CONDICIONES DE LAS EMOCIONES	11
TABLA 2: COMPARACIÓN DE LOS TRABAJOS EXISTENTES EN MER	14
TABLA 3: ADJETIVOS UTILIZADOS EN LA CLASIFICACIÓN MIREX.....	20
TABLA 4:PARÁMETROS ACÚSTICOS COMUNES PARA CLASIFICACIÓN DE EMOCIONES	22

Agradecimientos

Antes de presentar este proyecto me gustaría agradecer a mi director, José Ramón Beltrán Blázquez el haber confiado en mí para la realización de un proyecto tan ambicioso con el cual me he llegado a sentir especialmente motivado e interesado. Él me ha permitido realizar un Trabajo de Fin de Grado basado en la música, que era algo que no habría imaginado que pudiese llegar a conseguir. Me gustaría agradecerle todo el apoyo y las orientaciones que me ha ido dando para que este proyecto, que fue difuso al principio, fuera cogiendo forma hasta convertirse en lo que es hoy.

También me gustaría agradecer a mi codirectora, Piedad Garrido Picazo por la asistencia prestada, por hacer de madre académica en los momentos complicados y por toda la atención que le he robado a lo largo de este proyecto, sin la cual probablemente no podría haber salido adelante.

Por supuesto, agradecer a la Escuela Universitaria Politécnica de Teruel como institución por la oportunidad de realizar un proyecto alternativo e interesante sin problemas, además de por toda la formación y la oportunidad haber podido llegar hasta aquí estudiando lo que yo quería en mi ciudad.

Y cómo no, a la Universidad de Zaragoza al completo por estos años tan maravillosos en los que hemos peleado, hemos trabajado duramente, hemos fracasado, pero nos hemos vuelto a levantar, hemos viajado, hemos conocido gente, incluso hemos conocido el mundo laboral, pero sobre todo hemos aprendido. Aprendido cómo ser mejores estudiantes, cómo ser mejores profesionales. Aprendido a ser mejores personas.

1- Introducción

El trabajo aquí enunciado tiene como finalidad el estudio y desarrollo de un sistema sin supervisión que permita valorar determinadas canciones en función de su categoría emocional.

Tradicionalmente, el estudio de las emociones implícitas e inducidas por la música ha tenido siempre multitud de problemas de cara a su estudio y análisis, e incluso de cara a su validez y utilidad real. Hoy en día, en el marco tecnológico en el cual nos movemos, y existiendo enormes cantidades de datos circulando por las redes que crecen a cada instante, surgen nuevas necesidades de clasificación y selección, abriendo nuevas vías de investigación.

Este TFG busca no sólo realizar una investigación al respecto sino también diseñar e implementar un sistema sin supervisión que sea capaz de valorar y clasificar música según este criterio. Para conseguir cuantificar una característica que de por sí ya es muy abstracta, nos apoyaremos en estudios previos del este sector para alcanzar un modelo emocional adecuado.

El procedimiento de trabajo llevado a cabo ha sido el siguiente:

- **Acceso a información multimedia** de varios repositorios Web (Last.fm¹, Spotify)² de los cuales se extraen datos concretos sobre el archivo a analizar y conjuntos de etiquetas o *tags* creados por los usuarios. Para ello, se interactuará con las APIs de dichas plataformas mediante llamadas REST desde nuestro sistema.
- **Análisis y clasificación** de las canciones seleccionadas a partir de los datos que se obtengan, el sistema llevará a cabo labores de parseo y validación de información útil para eliminar cualquier dato que no sea relevante a nuestros fines.
- **Validación** mediante técnicas de minería de datos para verificar la correlación de los atributos de la canción y obtener un resultado válido.

¹ <https://www.last.fm/es/>

² <https://www.spotify.com/es/>

2- Estado del Arte

La música ocupa un espacio gigantesco en la vida cotidiana, y es debido a la capacidad que tiene de recompensar induciendo emociones a sus oyentes.

No hay una taxonomía de emociones generadas por la música hoy en día, por tanto, se deben utilizar modelos y medidas de áreas de investigación ajenas a la música. Esto conduce a una situación en la que no hay consenso sobre el proceso en el cual un estímulo musical se convierte en una emoción.

El estado del arte busca encontrar qué existe en la actualidad publicado sobre los siguientes interrogantes:

- ¿Qué estados emotivos son más frecuentemente inducidos por la música?
- ¿Son emociones específicas?
- ¿Cómo podemos clasificarlas y medirlas?
- ¿Difieren lo suficiente de las emociones habituales como para requerir una clasificación específica?

Para profundizar un poco más en el tema, se procederá a estudiar y comparar una serie de estudios realizados anteriormente y que abordan estos interrogantes.

2.1- Emociones evocadas por la música: Caracterización, clasificación y medida

El estudio aborda los siguientes interrogantes: ¿Qué estados emotivos son más frecuentemente inducidos por la música?, ¿Son emociones específicas?, ¿Cómo podemos clasificarlas y medirlas? Y ¿Difieren lo suficiente de las emociones habituales como para requerir una clasificación concreta?

2.1.1- El modelo de emociones actual

Parece que la música es capaz de inducir muchas más emociones de lo que simples etiquetas como: *alegría, tristeza, rabia, miedo, vergüenza, culpa...* serían capaces de representar. Aun así, en ausencia de una clasificación más adecuada, es entendible la utilización de dichas etiquetas.

Las emociones básicas aquí mencionadas no son muy adecuadas para esta medición ya que provienen de condiciones del entorno que afectan al individuo, y por ello, son de alta intensidad con el objetivo de preparar acciones en consecuencia. Las emociones generadas por la música no tienen efecto en el bienestar del individuo y no suelen ir seguidas de acciones [Zentner, Grandjean y Scherer 2008].

2.1.2- Investigación previa en emociones musicales

En 1912 se realizó uno de los primeros intentos para caracterizar emociones inducidas musicalmente de forma empírica. Ocho estudiantes de música como oyentes escucharon veinticuatro piezas, de 3'5 a 4'5 minutos. Las emociones reportadas fueron tan heterogéneas que se tiró la toalla. [Weld 1912]

En 1936 se desarrolló el círculo adjetival [Hevner 1936], que consistía en ocho grupos de emociones relacionadas, y que, según la posición en el reloj, marcaban lo lejos que una emoción estaba de otra. Hubo discrepancias en las décadas siguientes pero el modelo fue ampliamente utilizado.

En 1972 se propuso cuantificar las emociones mediante tres factores bipolares [Wedin 1972]:

- I. Alegría frente a tristeza,
- II. Tensión contra relajación y
- III. Solemnidad contra trivialidad

Tras esto se concluyó que las emociones musicales podían ser descritas a lo largo de 9 dimensiones de afecto.

Se investigó mucho en este ámbito, no obstante, dichos estudios no tuvieron un efecto notable en las prácticas habituales de investigación. En primer lugar, porque las emociones obtenidas no son las que experimentaban los oyentes, sino las inherentes a la música. Esta teoría se respalda cuando se confirma que el escuchar una música triste puede tener una experiencia positiva en una persona. En segundo lugar, porque idealmente estos estudios deberían haber tenido una compilación de términos afectivos que pueden ser experimentados escuchando música como base, y no suelen estar basados en teorías sino en experiencias personales. Un tercer problema de los estudios anteriores son los procedimientos usados para procesar y validar los datos, como utilizar muestras de población muy selecta e insuficiente, o valorar un resultado de manera puramente subjetiva. El cuarto y probablemente más notable problema de dichas investigaciones fue la ausencia de interés en buscar la existencia de emociones meramente relacionadas con la práctica de escuchar música y, en lugar de eso, compararlas con emociones rutinarias no musicales. Como consecuencia la noción de que las emociones musicales podrían ser de otro tipo que las “normales” permanece sin ser probada [Zentner, Grandjean y Scherer 2008].

2.1.3- Estudios previos interesantes:

Zentner, Grandjean y Scherer llevaron a cabo cuatro estudios muy interesantes que se van a explicar con más detalle en este apartado. Los dos primeros para compilar una lista de emociones musicalmente relevantes y para estudiar la distinción entre emociones sentidas y percibidas en grupos con intereses musicales diferentes. El segundo también examina la diferencia entre la experiencia de una emoción en una circunstancia habitual y en la música, y de los resultados se derivan términos para definir emociones experimentadas en distintos géneros musicales. En el tercero se busca extender el anterior usando una muestra de oyentes más grande y captando la emoción en el momento de ser evocada, y por último examinar la diferenciación de las emociones basándose en procedimientos *confirmatory factor analytic (CFA)*³. En el cuarto se intenta replicar el resultado del estudio tres con muestras diferentes y a la vez trata de validar el sistema actual de categorización de emociones.

³ Se trata de un método estadístico multivariable usado para probar lo bien que unas variables medidas representan una determinada muestra.

- **Estudio 1**

El objetivo era crear una lista comprensiva de palabras adecuadas para expresar emociones sentidas o percibidas. La muestra fue de noventa y dos estudiantes con una media de edad de veinticuatro años.

Se usó una lista de quinientos quince términos como posibles candidatos a descriptores verbales. La prueba consistía en hacer que los participantes dijeran si identificaban la esencia de un término, y si la emplearían para representar un estado emotivo en un formato binario (sí/no). Tras un proceso eliminatorio y selectivo en los resultados, quedaron ciento cuarenta y seis términos para el estudio dos.

- **Estudio 2**

El objetivo de este estudio es examinar cuales de los términos obtenidos en el estudio anterior son realmente relevantes en relación a la música. Se examina la frecuencia de dichos términos en contextos musicales y extramusicales, y como no es posible estudiar las respuestas emocionales a la música, se apoya en preguntar a los participantes qué es lo que sienten al escuchar determinados géneros.

Se usan doscientos sesenta y dos estudiantes, y se agrupan las preferencias musicales en cinco grupos (Clásica, Jazz, Pop/Rock, Latina, Tecno). Se obliga a que los sujetos elijan un estilo favorito.

Se facilitó un cuestionario a los participantes donde aparecía la lista con los ciento cuarenta y seis términos y una escala de cuatro puntos (nunca, a veces, a menudo, frecuentemente). Se dividió en dos partes, en la primera se tenía que indicar la frecuencia con la que esas emociones aparecían en los individuos, y en la segunda se tenía que indicar la frecuencia con la que los individuos identificaban esas emociones musicalmente. Al acabar se repartió otro formulario para hacer lo mismo, pero en situaciones ajenas a la música.

En los resultados, sesenta y cinco emociones fueron consideradas no musicales, de las cuales se decidió retener ocho concretas dado que consideraron prematuro eliminarlas, ya que eran términos de gran importancia, así que la muestra se quedó con ochenta y nueve términos emocionales.

Dichos factores fueron sometidos a análisis usando el método cuantitativo.

Factor	Adjective markers	α		
		Perceived	Felt	Everyday
1. Tender Longing	Affectionate, softened up, melancholic, nostalgic, dreamy, sentimental	.84	.82	.68
2. Amazement	Amazed, admiring, fascinated, impressed, goose bumps, thrills	.75	.81	.74
3. Tranquility	Soothed, calm, in peace, meditative, serene	.82	.84	.71
4. Joy	Joyful, happy, radiant, elated, content	.87	.80	.87
5. Activation	Disinhibited, excited, active, agitated, energetic, fiery	.79	.84	.75
6. Power	Heroic, triumphant, proud, strong	.79	.78	.71
7. Sensuality	Sensual, desirous, languorous, aroused (sexually)	.71	.78	.65
8. Transcendence	Ecstatic, spiritual feeling, mystical feeling, illuminated	.73	.75	.69
9. Dysphoria	Anxious, anguished, frightened, angry, irritated, nervous, revolted, tense	.85	.83	.82
10. Sadness	Sorrowful, depressed, sad	.79	.74	.76

Tabla 1: Factores extraídos, los adjetivos más discriminatorios y las Alfas de Cronbach para las tres condiciones de las emociones

Se compararon los resultados obtenidos con otros estudios diferentes, con las mismas bases, sobre otras muestras de población con una fuerte correlación, validando los resultados como se muestra en la tabla 1.

Se examinó hasta qué punto afectaba si una emoción era más percibida o sentida, y se concluyó que las emociones en general eran más inherentes a la música que inducidas en el individuo, especialmente en el caso de las escalas negativas como tristeza.

Respecto a la diferencia entre emociones musicales y extramusicales, aquí difieren mucho unas de otras, y difiere mucho en función de los grupos emocionales, en muchos de ellos incluso, se muestra predominante la emoción musical, a la vez que apenas se pueden encontrar en la vida cotidiana.

Las ratios también cambian mucho entre géneros musicales, en términos generales las emociones reflexivas se encuentran más entre el jazz y el clásico, las emociones enérgicas entre la música latina y el tecno, y las que sugieren rebeldía entre los oyentes del pop/rock.

Generalmente hablando, las emociones son menos sentidas en respuesta a la música que encontradas inherentes a ella. Particularmente, en las emociones negativas como tristeza o miedo podemos confirmar este supuesto ya que dichas emociones negativas no generan un estado similar en el oyente, dado que en la mayoría de contextos no se encuentra en peligro o bajo amenaza.

El principal problema de este estudio, es que los sujetos han sido consultados en un escenario imaginario en el que contestaban lo que opinaban en lugar de recoger las muestras directamente de un oyente en tiempo real.

• Estudio 3

El objetivo es extender los resultados del estudio dos con muestras más grandes y recogidas en directo, y además examinar si los estados emocionales inducidos pueden diferenciarse en unidades más grandes que pudieran servir como clasificación para las emociones evocadas por la música.

Los participantes se reclutaron de una serie de conciertos, unos dos mil participaron y ochocientos uno rellenaron el cuestionario. Redujeron las emociones a sesenta y seis gracias a unos estudios intermedios sobre la similitud de ciertos términos utilizados. La prueba consistía en valorar las emociones identificadas en determinados momentos de las obras con una escala del cero al dos.

Como resultado, de los ochocientos un cuestionarios la inmensa mayoría pertenecían a música clásica, y se eliminaron las emociones que obtuvieron menos de un cinco por ciento en la práctica real salvo la aflicción. También se aceptaron libres propuestas repetidas elegidas por los oyentes, éstas son admiración, fascinación, abrumación y tearful (al borde de las lágrimas).

Al final, los resultados demostraron que la muestra de emociones escogidas para el estudio dos era igualmente válida para conjuntos diferentes de población y para música en vivo, y, además, todas las observaciones tendían a converger con los resultados obtenidos en el estudio dos y en estudios realizados de forma similar en otros países.

Y, por último, la evidencia actual muestra que es posible diferenciar empíricamente estas emociones en nueve grupos (Wonder, transcendence, tenderness, nostalgia, peacefulness, power, joyful activation, tension and sadness), y éstas mismas pueden agruparse en 3 bloques (Sublimity, vitality and unease).

- **Estudio 4**

En este estudio se busca replicar la estructura en nueve dimensiones obtenida en los estudios anteriores con una nueva muestra de usuarios, con una selección exclusiva de fragmentos musicales no líricos y usando una escala optimizada incluyendo los términos añadidos en el estudio tres.

El principal objetivo de este último estudio es examinar la validez diferencial del modelo actual contra dos grandes modelos, el modelo de emociones discretas y el modelo emocional dimensional.

Esta prueba buscará sobre todo determinar si el modelo musical emocional es preferido sobre los otros dos por los oyentes, tanto por ajustarse más a las percepciones del oyente, como por discriminar mejor los diferentes tipos de experiencia emocional musical.

Los participantes son: miembros de coros profesionales, gente que leyó el artículo anterior y gente que leyó la solicitud respecto a este estudio, un total de doscientos treinta y ocho con cuarenta y siete años de media. Todos ellos afirmaron ser oyentes de música clásica.

Se trabaja con un formulario que representa los tres modelos emocionales. Ese cuestionario está formado por tres subcuestionarios en los cuales se muestra cada modelo y se valora cada emoción de uno a cinco.

El estímulo son dieciséis extractos de música clásica no vocal teniendo en cuenta que ya fueron utilizados en previos estudios de inducción de emociones, además fueron revisados por compositores, teóricos y filósofos y aprobados para esta prueba. La prueba se llevó a cabo en recinto cerrado mediante reproducción electrónica.

Los oyentes debían escuchar atentamente y con los ojos cerrados la primera reproducción del fragmento, y después de una breve pausa volver a escuchar el fragmento valorándolo. Diferentes grupos tuvieron diferente orden de presentación de los fragmentos.

Los participantes debían contestar en los tres modelos, después debían indicar cuál de los modelos consideraban más adecuado para capturar sus sentimientos inducidos por la música. También se formuló la posibilidad de realizar una continuidad del trabajo en casa con obras de libre elección.

La prueba mostró que la mayor parte de los oyentes encontraban más apropiado para describir su experiencia el modelo emocional musical, y que la representaba de manera más precisa.

Otro aspecto importante de cara a validar los resultados fue el acuerdo entre los oyentes de cara a la selección de una emoción específica. Tras realizar los análisis pertinentes también se llegó a la conclusión de que los oyentes coincidían más en las emociones presentadas en el modelo aquí presentado, seguido del modelo dimensional.

2.1.4- Valoraciones

Atendiendo a estos resultados, el modelo emocional musical específico superó en rendimiento a los otros dos. De todas formas, aun con este modelo, los resultados estuvieron lejos de resultar perfectos. Entonces, ¿podemos generalizar la relativa superioridad de este modelo para todas las emociones musicales?

Este estudio también arroja luz sobre el interrogante de si las emociones percibidas por la música son realmente emociones verdaderas, teoría ampliamente aceptada pero que en este estudio se pone en entredicho. En caso de que así fuera, se deberían haber obtenido unos resultados similares en los matices “emoción percibida” y “emoción inherente”.

En cualquier caso, sobre si son o no emociones la respuesta es complicada ya que no existe un acuerdo sobre lo que es una emoción.

“Es posible que las emociones sentidas en respuesta a la música sean simplemente un ejemplo de una categoría mucho más grande de emociones relacionadas con la evaluación de objetos cotidianos, situaciones y experiencias, los cuales, en esencia, ocupan un gran espacio en la vida de las personas.

Abriendo un camino en un área tan poco investigada, esperamos haber provisto un punto de partida a partir del cual los investigadores del día de mañana puedan alcanzar un entendimiento más profundo de las emociones musicales”
[Zentner Y Grandjean 2008]

2.2- Reconocimiento automático de la emoción musical

Dado el aumento que ha sufrido en los últimos años el uso de música para cualquier faceta de la vida, las metodologías para gestionar música basándose en metadatos no son suficientes. La manera en que se organiza y se recupera la música debe evolucionar para adaptarse a la creciente demanda de nuevas formas de acceso.

En base a esto, un sector que ha recibido mucha atención últimamente es el de la organización y la recuperación en base a emociones. Según Last.fm, la etiqueta emocional es la tercera más común.

Hacer que los ordenadores sean capaces de reconocer la emoción de la música también mejora la manera en que se relaciona con el usuario. Es posible reproducir música que se adapte al humor del oyente de diversas maneras [Yang, Chen 2012].

“En estudios psicológicos, las emociones suelen dividirse en tres categorías: emoción expresada, emoción percibida y emoción sentida o evocada [Gabrielsson 2002]”

Se ha trabajado mucho en los MIR (en inglés Music Information Retrieval) para reconocer la emoción percibida en la música. Se utilizan clases categóricas y se aplican técnicas de aprendizaje automático para entrenar clasificadores. Se utilizan análisis subjetivos para establecer un “suelo de verdad” para entrenar esas muestras.

Sigue sin haber consenso en la concepción de la emoción y su misma taxonomía. Existen tanto modelos continuos como conceptualizados. Principalmente se tienen en cuenta tres modelos recientes en reconocimiento de emoción musical, como se muestra en la tabla 2.

Methodology	Emotion conceptualization	Description
Categorical MER	Categorical	Predicting the <i>discrete emotion labels</i> of music pieces [Hu et al. 2008; Lu et al. 2006]
Dimensional MER	Dimensional	Predicting the <i>numerical emotion values</i> of music pieces [Eerola et al. 2009; Yang et al. 2008]
MEVD	Dimensional	Predicting the <i>continuous emotion variation</i> within a music piece [Korhonen et al. 2006; Schmidt et al. 2010]

Tabla 2: Comparación de los trabajos existentes en MER

2.2.1- Conceptualización de la emoción

De estudios empíricos se han sacado muchos modelos, y la mayoría pertenecen a un enfoque categórico o dimensional.

- Conceptualización categórica

Es esencial para esta aproximación el concepto de que las emociones primarias son limitadas y de ellas se derivan otras más específicas. El modelo más utilizado se compone de ocho grupos numerados con una serie de emociones dentro de ellos, grupos próximos comparten más características. El mayor problema es que los grupos son muy pequeños en comparación con la riqueza de la emoción percibida por la música, pero utilizar una granularidad más fina no soluciona el problema debido a la inherente ambigüedad de los términos utilizados, es más, podría saturar a los sujetos de pruebas, lo que lo hace inviable de cara a futuros análisis psicológicos.

- Conceptualización dimensional

El modelo dimensional se centra en el uso de dos ejes que simulan representaciones humanas de la emoción, que son la valencia y la actividad/estimulación. Una de sus fortalezas es que es un modelo muy simple pero potente y ofrece diferencias notables entre cada emoción de forma visual sencilla.

Aun así, tiene limitaciones, por lo que incluso se han llegado a proponer modelos tridimensionales añadiendo la potencia. No obstante, dicho modelo se presenta muy difícil de visualizar.

2.2.2- Rasgos musicales

Diferentes percepciones emotivas normalmente están asociadas con diferentes patrones de reglas acústicas. Normalmente la estimulación suele venir del *tempo*⁴, del volumen, y del timbre, mientras que la valencia⁵ suele venir del modo musical y de la armonía.

- **Energía:** La energía de una canción suele estar fuertemente relacionada con la excitación percibida [Gabrielsson y Lindström 2001]
- **Ritmo:** Normalmente un tempo rápido suele ser percibido como una alta actividad. Un tempo fluido suele conllevar valencia positiva, mientras que un tempo rígido y firme suele tener valencia negativa. La variabilidad del ritmo también es un aspecto importante
- **Melodía:** Se puede determinar el modo en el que está un fragmento musical, de modo que sopesando la cantidad de tiempo que se pasa en cada uno se puede obtener el modo principal
- **Timbre:** El timbre es más complicado de detectar, no obstante, se ha progresado mucho en análisis de ondas y espectros que permiten identificar instrumentos emisores de sonido diferentes. Diferentes instrumentos tienden a ocasionar diferentes reacciones en el usuario.

⁴ Tempo es un término italiano ampliamente utilizado en música que define la velocidad del ritmo de una pieza musical. Se suele medir en BPM (Beats per Minute)

⁵ Valencia en este estudio hace referencia a la positividad o negatividad de una determinada pieza musical

2.2.3- Reconocimiento categórico

Se han realizado numerosos esfuerzos para conseguir sistemas MER (Music Emotion Recognition) automatizados, y el tipo de música bajo estudio ha ido cambiando de música simbólica, a simples señales de audio y posteriormente a música clásica occidental. En la actualidad, suele elegirse la música clásica occidental ya que existe multitud de literatura en musicología y psicología de la música clásica.

Este enfoque categoriza piezas musicales por clases emocionales. Esto lo hace muy sencillo de manejar ya que es como una etiqueta más, de la misma manera que lo es el género y la instrumentación. Muchos estudios han entrenado modelos para predecir el contenido afectivo que más se ajusta a la pieza. El mayor problema es, como ya se ha expuesto antes, lo limitadas que son esas emociones.

2.2.4- Colección de datos

En ausencia de una base de datos general, algunos trabajos compilan la suya propia. Hay muchos problemas que impiden la realización libre de cualquier tipo de estudio como el copyright o la falta de consenso en la representación de las emociones.

En respuesta a la necesidad de una libre distribución de la música, MIREX (en inglés Music Information Retrieval Evaluation eXchange) y AMC (Audio Mood Classification) promueven investigación en MER y proporcionan comparaciones, ponen además los archivos de audio a disposición de los participantes. De todas formas, MIREX utiliza un modelo de cinco clústeres que no es muy utilizado en otros trabajos de MER. La comparación entre distintos modelos es virtualmente imposible. El problema de la taxonomía emocional sigue sin resolverse.

2.2.5- Preprocesamiento de datos

Las canciones objeto de estudio, normalmente deben preprocesarse y seleccionar solo entre unos veinte a treinta segundos de música que sean lo suficientemente característicos como para englobar la emoción de la canción.

2.2.6- Anotación subjetiva

Debido a que la emoción es subjetiva, para conseguir un suelo de verdad hay que ser muy cuidadoso. Los métodos de anotación suelen ser agrupados en “basados en expertos” y en “basados en sujetos”. El método basado en expertos requiere unos pocos expertos musicales que anoten, si es imposible llegar a un acuerdo en alguna pieza, se descarta. El método sujeto, suele pedir opiniones a sujetos no entrenados y luego se median los resultados. Se deben seguir medidas para evitar la fatiga humana.

Dado que la percepción musical es multidimensional, ¿deberíamos preguntar a los oyentes si les suena la canción?, ¿ocultarles los líricos o mostrarlos en otros idiomas? De momento no hay consenso.

Para mitigar la dificultad de la anotación subjetiva, una tendencia habitual es obtener las etiquetas emocionales de sitios como AMG y Last.fm, que se puede hacer con un simple script. El problema es que la calidad es inferior respecto a la obtenida mediante anotaciones subjetivas.

La última tendencia es efectuar juegos en los que la gente compita directamente contra otro jugador. En estos juegos se busca contribuir emocionalmente haciendo que los jugadores valoren una pieza y obtengan más puntos por acercarse más a los resultados generales de los demás jugadores. Normalmente se obtienen anotaciones gratuitas de buena calidad.

2.2.7- Entrenamiento del modelo

Después de obtener un suelo de verdad lo que hay que hacer es entrenar un sistema de aprendizaje automático para aprender las relaciones entre emociones y rasgos musicales. Suelen emplearse técnicas de aprendizaje automático como redes neuronales, k-nearest, árboles de decisión, etc.

2.2.8- Reconocimiento dimensional

El modelo circunflejo se empezó a utilizar por investigadores del MER para ver la variación emocional en una pieza clásica. La idea de representar la emoción general como un punto en el plano, aparece en recientes estudios bajo el supuesto de que la emoción general de una canción popular implica menos cambio que una pieza clásica. Los autores marcaron MER como un problema de regresión. De esa forma, es posible marcar un punto en el plano para obtener una lista de canciones asociadas a esa emoción o trazar una línea para obtener una lista de canciones comprendidas en ese dominio.

2.2.9- Detección de la variación en la emoción musical

Un aspecto importante son las dinámicas temporales. Muchos investigadores se han centrado en fragmentos homogéneos, no obstante, muchos estilos muestran o evocan diferentes emociones conforme avanza el tiempo, por lo tanto, es importante valorar esa relación. En la ilustración 1 se muestran varios estudios de estas dinámicas.

El modelo categórico no se muestra del todo competitivo ya que requiere de una constante entrada de términos emocionales, por eso el modelo dimensional es más adecuado.

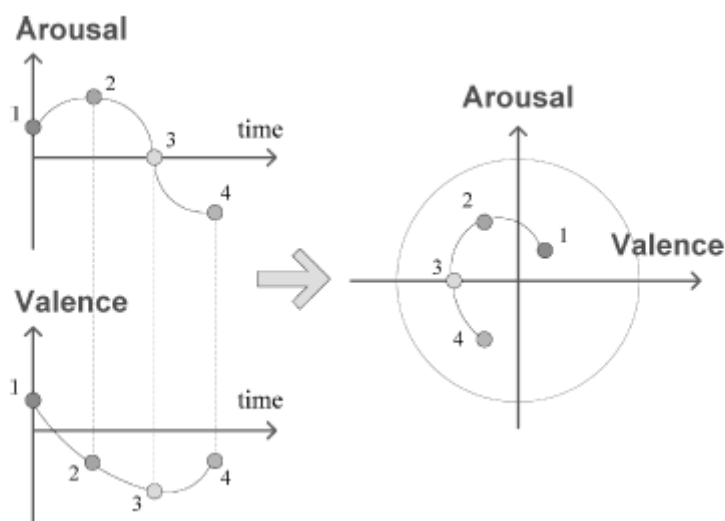


Ilustración 1: Muestra de variación en la detección de emociones. Combinando los valores en valencia y excitación podemos formar la curva afectiva (derecha), la cual representa los cambios dinámicos del contenido afectivo de una pieza musical

2.2.10- Desafíos

Debido a que el MER sigue en su infancia, tiene muchos problemas abiertos:

- Dificultad en la notación de la emoción

Para obtener el suelo de verdad necesario para entrenar modelos automáticos normalmente se realizan estudios subjetivos. Dado que el sistema MER está pensado para usarse en contextos habituales, la notación debería llevarse a cabo por gente común. El problema de obtener las etiquetas se ha aliviado recientemente gracias a sitios como AMG y Last.fm. No obstante, el proceso dimensional necesita valores numéricos que no pueden encontrarse en los repositorios en línea.

- Subjetividad en la percepción de la emoción

La percepción de emociones es intrínsecamente subjetiva y depende de muchos factores. Es muy difícil alcanzar un consenso de cuáles son los términos que mejor se ajustan a una pieza. Diferentes personas responden opiniones diferentes frente a la misma canción, como podemos observar en la ilustración 2:

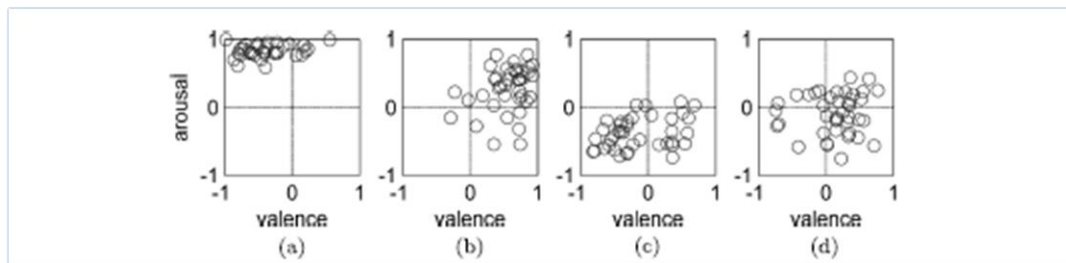


Ilustración 2: Anotaciones emocionales para 4 canciones: (a) *Smells Like Teen Spirit* de Nirvana, (b) *A Whole New World* de Peabo Bryson y Regina Belle, (c) *The Rose* de Janis Joplin, (d) *Tell Laura I Love Her* de Ritchie Valens. Nótese como la percepción es subjetiva.

En base a esto, algunos estudios proponen enfoques diferentes en los que miden, en porcentaje, la fuerza de cada emoción en las piezas musicales.

Para el enfoque dimensional del MER se necesitan metodologías diferentes ya que son expresadas por valores numéricos. Se plantean soluciones mediante dos mediciones diferentes, la predicción de la percepción general y la predicción de la diferencia entre el usuario y la percepción general. Estos métodos son viables, pero carecen de un entorno computacional sólido.

2.2.11- Hueco semántico entre el rasgo musical de bajo nivel y la emoción humana de alto nivel

La viabilidad de los sistemas MER recae en la precisión del reconocimiento emocional, pero es muy difícil predecir valores o etiquetas debido al propio conocimiento humano. No está claro qué elementos intrínsecos de la música generan qué estímulo. Consecuentemente el rendimiento de los sistemas basados en la explotación de recursos de bajo nivel alcanza un límite.

Para mejorar este límite, puede ayudar el empleo de técnicas adicionales, como la extracción de los líricos, las progresiones acórdicas o el género. Se ha demostrado empíricamente que incluir dichos aspectos mejora la precisión.

Por lo tanto, según lo aquí expuesto, a pesar de llevar un gran esfuerzo por detrás, los sistemas MER todavía son muy jóvenes y tienen muchos problemas sin resolver o incluso desconocidos.

2.2.12- Explotación del timbre vocal

El timbre de la voz cantante suele ir relacionado con la percepción emocional por lo que debería ser tomado en cuenta. Para la realización de esta medida sería clave poder anular la música acompañante, cosa posible salvo en los casos de los instrumentos que comparten la frecuencia de la voz cantada, para cuyos casos se requerirían acciones avanzadas, aún lejos de poderse hacer con éxito.

2.2.13- Conectar los sistemas MER dimensionales y categóricos

Ambos sistemas ofrecen ventajas complementarias, por lo que resultaría interesante conectarlos de cara a obtener un sistema más efectivo.

Debido a que los usuarios pueden sentirse poco identificados con la esencia de la valencia, se puede ayudar al modelo con emociones para guiar al usuario y aportar un mayor entendimiento del sistema bidimensional gracias al sistema categórico.

2.2.14- Considerar los factores situacionales

Nuestra respuesta a la música depende de una síntesis entre factores musicales, personales y situacionales. Sería genial que los MER pudieran **detectar el humor del oyente o el entorno** para modificar las predicciones **o incluso utilizar esa información para recomendar música de forma activa al oyente.**

2.3- Maneras de reconocer emociones en música

Normalmente la música es referida como el lenguaje de la emoción, no obstante, a la hora de desarrollar sistemas automáticos que clasifiquen la música en función de su contenido emocional nos encontramos con un problema que normalmente carece de una respuesta bien definida. Se vuelve particularmente importante el explorar un amplio rango de búsquedas en reconocimiento de emoción musical, particularmente haciendo hincapié en métodos que utilicen información contextual (como sitios web, etiquetas y líricos).

“Determinar el contenido emocional de la música es, por naturaleza, un esfuerzo multidisciplinar que implica no sólo procesado de señales y aprendizaje automático, sino que requiere también un entendimiento de percepción auditiva, psicología y teoría musical.

Pero la emoción no se encapsula únicamente en el audio (el contexto social, por ejemplo, juega un rol importante), de modo que vamos a estudiar enfoques incorporando metadatos, como etiquetas y letras de canciones” [Yang, Chen 2012].

2.3.1- Investigación psicológica

Muchos estudios, usando medidas categóricas o escalares/vectoriales, indican la importante distinción entre la percepción personal de las emociones expresadas por la música y las inducidas por la música [Yang, Chen 2012].

En este estudio se hará hincapié en sistemas que identifiquen la emoción expresada en lugar de la inducida por la música.

Como modelos de representar la música se presentan dos modelos categóricos, el de Zentner y la clasificación MIREX, la cual se presenta en la tabla 3:

Clusters	Mood Adjectives
Cluster 1	passionate, rousing, confident, boisterous, rowdy
Cluster 2	rollicking, cheerful, fun, sweet, amiable/good natured
Cluster 3	literate, poignant, wistful, bittersweet, autumnal, brooding
Cluster 4	humorous, silly, campy, quirky, whimsical, witty, wry
Cluster 5	aggressive, fiery, tense/anxious, intense, volatile, visceral

Tabla 3: Adjetivos utilizados en la clasificación MIREX

Otros estudios proponen que las emociones pueden ser escaladas y medidas de manera multidimensional. El modelo más aceptado es el modelo *valence-arousal* que mide en función de la intensidad y la polaridad de la emoción. Dicho modelo está presentado en la figura 3.

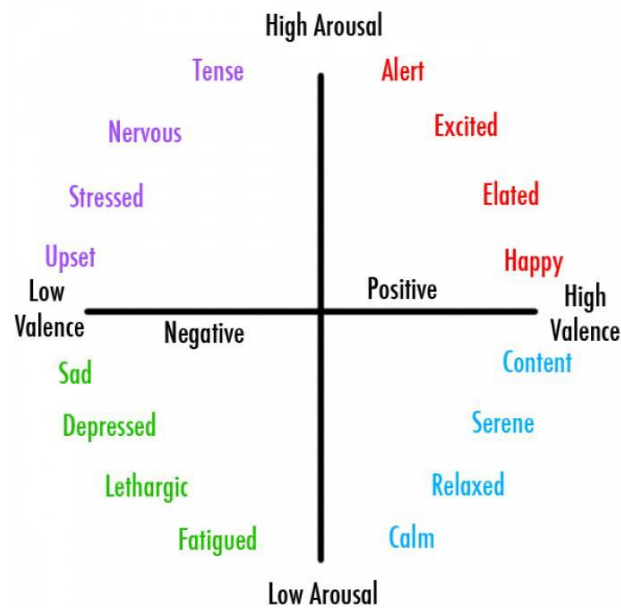


Ilustración 3: Clasificación según el modelo valence-arousal

2.3.2- Entorno de trabajo para reconocimiento de emociones

Se pueden estimar emociones de muchas maneras, se puede preguntar directamente a oyentes humanos, se puede utilizar minería de metadatos en textos como documentos web, biografías u opiniones, y también es posible valorar mediante el procesamiento de señal y el aprendizaje automático.

La anotación humana se puede realizar mediante encuestas personales, aunque experimenta problemas de cara a compartir los resultados y hacerlos visibles de forma sencilla. Un segundo enfoque es el social tagging por ejemplo de Last.fm, que permite a los usuarios contribuir poniendo tags a las piezas musicales, cosa que hacen posible gracias a sus APIs públicas, aun así, se enfrenta a múltiples problemas debido a la dispersión de resultados, errores de introducción, introducción maliciosa, etc... También se han utilizado juegos de diferente manera para alentar a los oyentes a que valoren emocionalmente una pieza. En la siguiente figura tenemos el juego Moodswings que funciona de dicha forma, los dos jugadores escuchan una pieza musical, y la valoran emocionalmente en el eje Valence-arousal, cuanto más se acercan el uno del otro, más puntuación reciben.

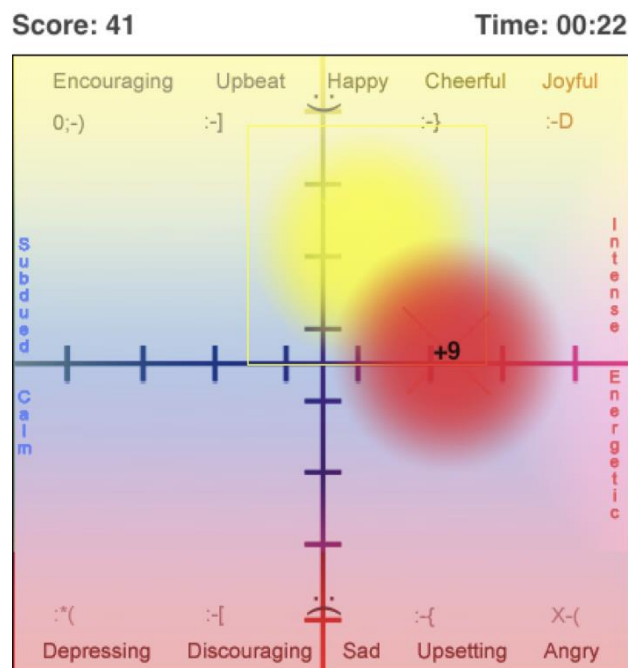


Ilustración 4: Moodswings, juego con propósito de ayudar al entendimiento de las emociones en la música

La información contextual puede obtenerse de documentos web como biografías, reviews de álbumes y de canciones. Hay múltiples sistemas de reconocimiento que encolan motores de búsqueda, blogs mp3, o sitios web. El problema de estas búsquedas es que suelen ser muy ruidosas, debido a mucha información irrelevante y datos inútiles.

Muchos sistemas de minería de datos utilizan documentos asociados a una canción y los vectorizan. Este espacio vectorial es muy útil para calcular similitud entre música. Recientemente Knees ha propuesto el **relevance scoring**, una alternativa prometedora.

Los social tags se han utilizado para cumplir tanto labores de búsqueda de música como de establecimiento de emociones. Un enfoque ha sido el de agrupar los tags en distintos modelos psicométricos en forma de descriptores, ya que hacer de cada tag una clase única genera una

situación inmanejable con demasiadas dimensiones y fracaso a la hora de tener en cuenta la similitud entre términos. Hay varias formas de agrupar los tags en grupos homogéneos, se puede hacer de forma manual pero no es escalable dada la magnitud de las muestras.

La extracción de emociones desde los líricos es particularmente difícil porque la extracción de rasgos y esquemas para etiquetar líricos **no son triviales**. Por eso se utilizan como herramienta de apoyo.

2.3.3- Análisis del contenido del audio

Claramente el propio audio puede derivar en una identificación de emociones intrínsecas, y dado lo rápido que se expanden las librerías musicales está claro que la anotación manual no escala eficientemente.

Las emociones pueden estar influenciadas por aspectos acústicos como el tempo, el timbre, la armonía y la carga sonora. Los rasgos más comúnmente estudiados se muestran en la tabla a continuación:

Type	Features
Dynamics	RMS energy
Timbre	MFCCs, spectral shape, spectral contrast
Harmony	Roughness, harmonic change, key clarity, majorness
Register	Chromagram, chroma centroid and deviation
Rhythm	Rhythm strength, regularity, tempo, beat histograms
Articulation	Event density, attack slope, attack time

Tabla 4: Parámetros acústicos comunes para clasificación de emociones

También se han estudiado múltiples características sonoras como el sonograma, el espectrograma, patrón de fluctuación... Los mejores resultados se obtenían al usar el máximo posible de ellos.

Por otro lado, el análisis de audio basado en su contenido puede usar tanto un modelo categórico como paramétrico.

2.3.4- Combinando múltiples análisis distintos

Está claro que algunos aspectos de la información musical no están implícitos en el audio y que usando únicamente rasgos acústicos existe un techo de rendimiento que no puede mejorarse sin recurrir a otras fuentes. Esto lleva a un aumento de investigación relacionada con la inclusión de sistemas múltiples para mejorar el reconocimiento en sistemas Music-IR (music information retrieval).

Audio y líricos

Diversos géneros musicales (como música navideña) son más fáciles de detectar usando el texto. Esto motivó el desarrollo de clasificaciones utilizando múltiples combinaciones de recursos auditivos y textuales.

Muchos estudios demuestran que la utilización de audio y líricos combinados mejoran el rendimiento general en todas las categorías. [Meyers 2007]

Se han realizado estudios recientemente implicando parámetros, líricos y metadatos para automáticamente etiquetar música en una versión discreta del espacio V-A. El mejor sistema utilizando estas tecnologías consiguió una precisión por valor del 64.1% en valencia y un 60.9% en excitación [Schuller, Dorfner, Rigoll 2010].

Audio y tags

También se ha hecho mucho hincapié en la incorporación de tags y rasgos musicales de bajo nivel. Se recogían los tags de Last.fm y los rasgos musicales, y después se combinaban, se agrupaban en uno de los clusters MIREX y también en uno de los cuadrantes del espacio V-A. En cualquier caso, se pudo observar que el trabajo conjunto de los dos enfoques y la compensación mejoraba la precisión global del sistema.

Como conclusión de los estudios abordados, identificar emociones musicales sigue siendo un desafío principalmente debido a las inherentes ambigüedades de las emociones humanas. El mayor avance ha sido posible a partir del uso de anotaciones y rasgos basados en el contenido. A pesar de todo, el sistema más preciso está muy lejos de poder realizar predicciones. No obstante, debido a que los avances se han dado combinando varios objetos de estudio con la posibilidad de añadir elementos sociales, está claro que la línea de investigación deberá continuar explotando las redes y realizando procesos de extracción de información y almacenes de datos.

2.4- Software y tecnologías trabajando hoy en día en este ámbito

Existen multitud de desarrollos que buscan arrojar luz sobre este aspecto tan complejo de la música. Hoy en día, múltiples aplicaciones continúan realizando este tipo de trabajo, algunos en forma de juegos con propósito, como Moodswings, otros juegos con el mismo objetivo como MajorMiner, o Herd It, grandes aplicaciones como Spotify, iTunes o Last.fm también están destinando cada vez más esfuerzos en continuar esta línea.

Yi Hsuan-Yang, una celebridad en este campo y referenciado numerosas veces a lo largo de este estudio, incluso tiene su propia aplicación elaborada a partir de los resultados de sus estudios. Su aplicación, denominada “Mr. Emo” consiste en un espacio dividido por los ejes valence-arousal, en los cuales, seleccionando sobre una zona concreta el sistema devuelve un listado de canciones asociadas a ese estado anímico.⁶

Las grandes aplicaciones de reproducción, como Last.fm, Itunes, Youtube o Spotify también se ponen constantemente al día y buscan nuevas formas de etiquetar el contenido para acceder a él mediante diferentes criterios. El factor emocional ha sido últimamente un factor de peso, como podemos observar en la reciente incorporación de Playlists emocionales a las listas de Spotify, o el “social tagging” que Last.fm lleva años efectuando. Es una realidad, esta línea de investigación tiene futuro y las grandes empresas cada vez destinan más recursos a realizar grandes procesos de minería y almacenes de datos para traer más luz a este ámbito.

No obstante, sigue habiendo muchas cosas sin hacer, y es que, aunque en éstos últimos años el desarrollo se haya acelerado notablemente [Yang y Chen 2012], todavía quedan muchos aspectos sin explorar y nuevas vías por descubrir. Hay muchas combinaciones de parámetros que todavía no han sido explotadas y analizadas [Kim, Schmidt, Migneco, Richardson, Scott, Speck, Turnbull 2010].

Dejando la vía de investigación abierta, y planteando posibles nuevas rutas de desarrollo, este TFG encuentra su oportunidad de prosperar, retomando una de estas investigaciones y aportando nuevos resultados con muestras de datos que hasta el momento no hayan sido comparadas.

Desde luego, para este TFG, buscaremos continuar con esta línea de investigación utilizando los recursos que han dado buenos resultados previamente y explorando diferentes vías de trabajo, con la intención de traer algo más de luz a este sector todavía desconocido.

⁶ <https://www.youtube.com/watch?v=ra55xO20UHU>

3- Plan de trabajo

Teniendo en cuenta el estado del arte, y buscando un nicho de mercado para el TFG, se planteó la posibilidad de realizar un estudio en este mismo sector, utilizando unas fuentes diferentes y de forma automática. Un planteamiento distinto a partir de una base de verdad lo más confiable posible con nuevos enfoques.

Este sector es complicado, obtener bases sobre las que trabajar no es algo trivial y suele requerir reunir grandes muestras de gente y someterlas a experimentos largos y tediosos para al final no poder obtener enormes cantidades de datos, y ni aun así garantizar la estabilidad de las muestras de datos.

Las pruebas precedentes han demostrado que el experimento social de Last.fm, el social tagging, ha brindado buenos resultados en este sector, de tal manera que acceder a esos tags será nuestra principal prioridad. Dichos tags serán considerados nuestra base de verdad, y trabajaremos a partir de ellos.

Desde el principio ya se plantea un problema, la subjetividad del estudio. Los tags de Last.fm son elaborados por personas y por supuesto poseen una alta probabilidad de contener datos basura o irrelevantes. Por este motivo, se vuelve imprescindible diseñar algún mecanismo que nos permita seleccionar tan sólo los tags relevantes.

Una vez tenemos nuestra base de verdad necesitamos una muestra a la cual enfrentarla y contrastarla para intentar buscar relaciones. Se analizaron las principales redes sociales en un intento de poder extraer información útil de cara al estudio. Se planteó la posibilidad de analizar contenido de YouTube, Spotify, foros de opinión, comentarios en redes, pero, aunque ciertamente podría haber sido interesante, analizar grandes cantidades de texto en busca de datos relevantes es bastante complicado y muy difuso, gasta demasiados recursos y no brinda resultados concluyentes a los pequeños desarrolladores.

A partir de este punto hubo dudas sobre la continuidad de la propuesta y cómo sacarla adelante, ya que todas las posibilidades planteadas eran o poco precisas, o demasiado extensas para el TFG.

Finalmente, indagando por la API de Spotify se halló un método cuanto menos interesante. Una función que dada una canción permitía devolver una serie de parámetros físicos en tiempo real, de lo más diversos y con una alta implicación emocional.

Algunos de estos parámetros son los valores numéricos aproximados de la energía, el valor acústico, el volumen, el tempo, la valencia o el tono. Todos estos valores se calculan para cada canción introducida en Spotify de manera automática y permanecen guardados. Este hecho fue muy conveniente ya que nos permite obtener esos parámetros de cualquier canción que haya sido introducida a Spotify y procesada.

Se realizó una prueba para asegurar que toda canción existente en Spotify posee su análisis de propiedades. Se solicitó en masa obtener las características de una gran muestra de canciones aleatorias y todas ellas permitían obtener el resultado, así que decidí ir un poco más allá. Aprovechando que en el momento de la experimentación estaba grabando una canción con mi grupo musical, y sería subida a Spotify, probé si incluso temas subidos por artistas poco reconocidos eran analizados. Tardó unos 4 días en obtener los análisis, pero los obtuvo, incluso teniendo menos de cien reproducciones, por lo que podemos asumir que toda canción que se agregue a la plataforma es analizada a corto plazo.

Tras esta prueba se confirmó que era posible afrontar este reto, que era posible intentar buscar relaciones entre las etiquetas escritas por oyentes de canciones y los parámetros físicos que ayudan a definir una canción. Con esta premisa, la realización de nuestro Sistema de Música y Emociones (de ahora en adelante SME) se llevó a cabo.

4- Análisis

Dada la naturaleza del proyecto, se requiere de una aplicación que realice las siguientes tareas:

- Obtención de datos de las fuentes
- Conversión de los datos obtenidos en datos manejables por el sistema
- Análisis lingüístico valorativo
- Cruce y asociación de datos

Para alcanzar tales metas, SME deberá cumplir una serie de requisitos, en torno a los cuales girará la elección de tecnologías.

4.1- Requisitos

- El sistema deberá conectarse en tiempo real a las APIs de Last.fm y Spotify
- El sistema debe permitir una forma de especificar el objeto de búsqueda
- El sistema deberá advertir en caso de que no se pueda efectuar la búsqueda
- El sistema contará con una base de datos relacional para almacenar los resultados de las búsquedas
- Los datos obtenidos de las búsquedas deberán ser procesados y convertidos en datos utilizables por el sistema si es necesario
- El sistema debe permitir visualizar el contenido de la base de datos actual
- El sistema deberá permitir al usuario realizar un análisis de los datos obtenidos
- El sistema deberá aplicar los cálculos necesarios para poder cuantificar los datos obtenidos en el análisis
- El sistema deberá utilizar los datos cuantificados para procesar la información y dar resultados

En base a los requisitos básicos establecidos previamente se procede a investigar las tecnologías de libre distribución que se ajustan a las necesidades del proyecto y que nos permitirán alcanzar una solución.

4.2- Sistema Operativo

Todo el desarrollo se realizará en entornos virtuales, debido a que ofrecen ciertas ventajas, particularmente la posibilidad de capturar instantáneas de progreso y retornar rápidamente a puntos “seguros” cuando se realizan tareas críticas, o incluso realizar desarrollos paralelos.

La elección del Sistema Operativo (SO) se realizará dando por hecho que se va a trabajar en un entorno Linux, debido a la sencillez y agilidad que aporta a la hora de desarrollar y trabajar.

Considerando la experiencia previa del usuario en este entorno, y que no hay grandes diferencias entre distribuciones que afecten de manera directa al proyecto, se elegirá la distribución KUbuntu.

Esta decisión se apoya principalmente en la comodidad, la facilidad de uso, la ligereza que aporta la interfaz KDE (en inglés “Kool” Desktop Environment), el hecho de compartir los repositorios de Ubuntu, y la existencia de una comunidad enorme que siempre es de gran ayuda de cara a resolver los problemas que van surgiendo.

También se hace necesario utilizar ocasionalmente un entorno Windows, por motivos que veremos más adelante.

4.3- Software de virtualización

Debido a que se ha optado por trabajar sobre máquinas virtuales, se precisa un software de virtualización de libre distribución que nos permita trabajar cómodamente.

Se eligió la versión gratuita de VMware porque presenta unas características suficientes para el desarrollo completo del proyecto, adicionalmente durante el transcurso de la ingeniería fue el software utilizado cuando surgía una necesidad de virtualización por lo que la experiencia en ese entorno es superior por parte del estudiante.

4.4- Lenguaje de programación

La elección del lenguaje de programación es probablemente la más importante de cara a obtener el éxito en el proyecto debido a que será la base, planteará el enfoque y el paradigma a seguir y nos impondrá sus propias limitaciones de cara al desarrollo.

Para este proyecto no se requiere un enfoque orientado a objetos, que es el enfoque más trabajado durante la ingeniería, por lo que se deberán buscar nuevas alternativas más adecuadas a este fin y realizar una formación suficiente para ser capaz de realizar un desarrollo competente en un entorno nuevo.

Después de una investigación sobre el tema y los consejos de varios expertos en el tema, se decidió que el lenguaje de programación utilizado sería Python⁷. Esa decisión fue apoyada por la buena crítica que tiene Python en general por la comunidad de desarrolladores, además de por su sencillez y limpieza en el código.

Python es un lenguaje interpretado de alto nivel, muy sencillo de aprender y de leer, además cuenta con infinidad de librerías de apoyo en prácticamente todos los ámbitos. Para reforzar la decisión, suele ser el lenguaje utilizado en los desarrollos científicos por lo que hay una enorme comunidad muy activa trabajando constantemente con este lenguaje, y la mayor parte de interfaces están principalmente diseñadas para su acceso o instalación desde Python. Además, es un lenguaje multiparadigma que cubre las necesidades del proyecto ya que soporta orientación a objetos, programación imperativa y programación funcional.

⁷ <https://www.python.org/>

Además, para nuestro estudio estadístico y minería de datos se estudiará y se empleará el lenguaje R⁸, de análisis estadístico. R es un lenguaje muy potente que permite instalar cómodamente librerías de análisis y minería. Además, cuenta con el IDE RStudio, que usaremos para asistirnos en el proceso de aprendizaje del lenguaje nuevo. Este hecho es el que obliga a tener que trabajar en un entorno Windows debido a que la versión de RStudio para Linux no está funcionando en mi distribución.

4.5- Sistema gestor de base de datos

Debido a que nuestro sistema necesita almacenar una serie de datos de forma persistente se vuelve necesario contar con el apoyo de una base de datos.

Debido a que el problema no requiere de una base de datos compleja ni con fuertes dependencias, se elige una base de datos relacional sencilla y de libre distribución. Con estas características se eligió MySQL como sistema gestor debido a que se habilitó un espacio en un servidor de la Universidad de Zaragoza en el que trabajar, y el gestor utilizado era MySQL, el cual era perfectamente suficiente y sencillo de utilizar y conectar desde Python.

Más adelante se planteó la posibilidad de migrar el sistema a un NoSQL como MongoDB⁹, lo cual tendría varios efectos positivos. En primer lugar, sería exportable a JSON, lo que nos dota de una gran versatilidad, y, por otro lado, escala muy bien en horizontal, y debido a que obtenemos los parámetros de las canciones desde fuentes externas, cabe la posibilidad de que, por cambios inesperados, la tabla tenga que crecer en horizontal. Este problema quedaría fácilmente resuelto en un modelo NoSQL.

4.6- Fases

El desarrollo de este proyecto se dividirá en la consecución de tres fases fundamentales, que se corresponderán con el proceso evolutivo del prototipo final de alta fidelidad.

4.6.1- Primera fase: Acceso a datos multimedia

La primera fase estará enfocada sobre todo a crear una base estable para el desarrollo, iniciarse en el lenguaje y el entorno y comprobar que los intentos de conexión con los servidores externos se realizan correctamente.

En este punto se efectuarán los siguientes puntos de control:

- Prueba básica de funcionamiento del lenguaje
- Prueba de llamadas REST desde el lenguaje
- Prueba de compatibilidad con APIs de desarrollo
- Prueba de obtención de datos externos

⁸ <https://www.r-project.org/>

⁹ <https://www.mongodb.com/es>

4.6.2- Segunda fase: Almacenamiento y validación de los datos

Una vez que nuestro sistema ya es capaz de obtener datos de plataformas externas correctamente, es necesario almacenarlos correctamente, por tanto, esta segunda fase implementará la conexión con la base de datos, la cual deberá diseñarse en función de las necesidades del proyecto.

Adicionalmente, se diseñarán en este punto los algoritmos de obtención y almacenamiento automáticos, además de los procedimientos de pre-procesado de los datos.

De esta manera, se establecen los siguientes puntos de control en esta fase:

- Prueba de inserción básica en base de datos con datos originales
- Prueba de robustez insertando de manera violenta muchos datos
- Prueba de inserción de búsquedas personalizadas automáticas multi-idioma
- Elaboración de los filtros de inserción inteligente
- Prueba de integridad tras pre-procesar los datos

4.6.3- Tercera fase: Análisis y minería de datos

Por último, una vez ya dispongamos de una base de datos llena, nuestro último objetivo consiste en estudiar los parámetros de nuestros datos y realizar procesos estadísticos y minería para ser capaces de entender la muestra y poder sacar conclusiones.

Para ello se diseñará un pequeño sistema externo que seguirá los siguientes pasos:

- Conexión con la base de datos
- Obtención de las muestras como dataset que permita aplicarle minería
- Estudio y aplicación de funciones disponibles

Con este tercer y último prototipo completaremos el objetivo del proyecto y cerraremos el sistema de apoyo a la toma de decisiones.

5- Diseño

5.1- Interfaces

Nuestro sistema carecerá de interfaces con el usuario ya que no se ha visto necesaria una interacción con el usuario en tiempo de ejecución, sino que responderá como un conjunto de funciones invocadas según la necesidad del momento. La única respuesta que tendrá el usuario serán notificaciones del proceso de las operaciones por consola, y avisos en caso de errores o detenciones inesperadas.

Dichas funciones se invocarán en primer lugar desde la terminal del propio usuario, y, posteriormente, se desplegarán para su uso desde cualquier otro punto como servicio externo mediante la invocación de la función.

5.2- Despliegue

El despliegue del proyecto implicará el Sistema de Información (SI) local del estudiante, donde se desarrollará el proyecto y los avances, además de la base de datos habilitada en un servidor de la Universidad de Zaragoza (UZ). El sistema informático del estudiante se conectará con ese servidor cada vez que se necesite guardar o acceder a la información del sistema. Se ha preferido no tener la base de datos almacenada localmente en el sistema del usuario por razones de seguridad y de accesibilidad, ya que, de esta manera, cualquier sistema con los permisos necesarios puede acceder a los datos de la aplicación.

El diseño lógico de nuestro sistema se basaría en un fichero principal, el cual accedería a las diferentes funciones alojadas en ficheros diferentes o en el mismo. En cualquier caso, toda la lógica de la aplicación se encuentra almacenada en el mismo directorio salvo la parte de minería, que la consideraremos un elemento auxiliar.

5.3- Diagrama de actividad

De esa manera, nuestro sistema realizará una actividad principal, compuesta de muchas sub-tareas que se pueden resumir en el siguiente diagrama de actividad:

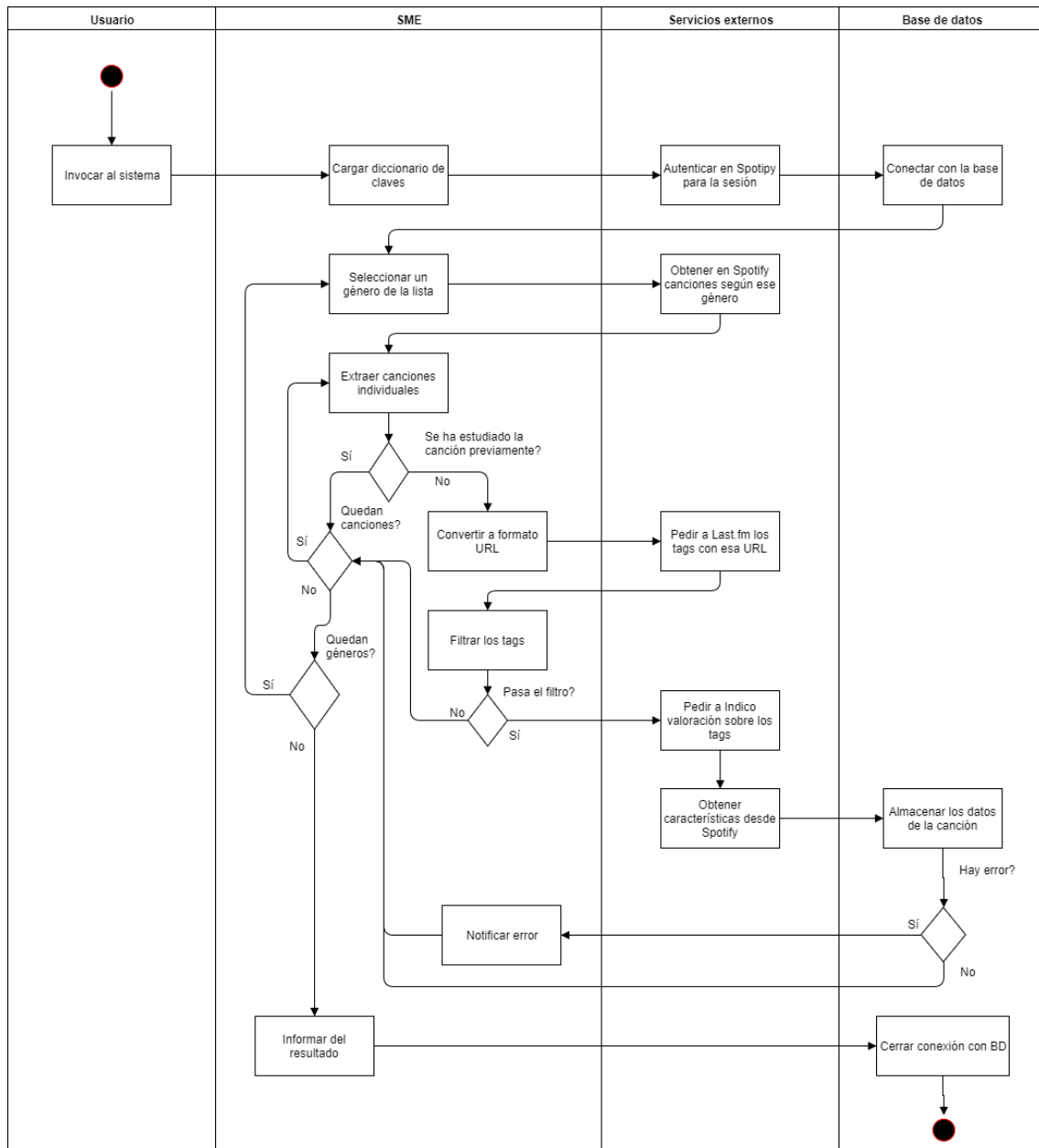


Ilustración 5: Diagrama de actividad de SME

Como se puede observar, la interacción del usuario a lo largo del proceso es nula, este es el principal motivo por el cual no denominamos a SME como una aplicación, sino como un sistema de apoyo. De hecho, este sistema idóneamente será embebido dentro de otro sistema que aprovecharía sus funcionalidades invocándolo directamente, eliminando por completo la presencia del usuario. Estas propiedades son las que convierten nuestro sistema en una “caja negra”.

5.4- Servicios externos

Los servicios externos proporcionados por Spotify, Last.fm e Indico son considerados actividades en sí y tienen su propio tratamiento independiente. Esto se diseña así para prevenir la posibilidad de que ante la caída permanente de una de las fuentes no haya que rediseñar toda la arquitectura, sino solo la parte afectada. El hecho de que recurramos a tantos servicios de proveedores que escapan a nuestro control nos obliga a llevar ciertas estructuras de control y prevención en caso de problemas.

No obstante, debido a la naturaleza del algoritmo y del sistema en sí, y dado que lo hemos construido en torno a la funcionalidad de Spotify que nos permite obtener canciones conocidas aleatorias según géneros, si el servicio que cesa su funcionalidad permanentemente es Spotify, nuestro diseño dejaría de ser válido y sería necesario un proceso de reingeniería, no siendo el caso frente a un cese de servicio por parte de Last.fm o Indico, los cuales solo obligarían a buscar otro proveedor y conectar con sus sistemas.

En cualquier caso, vamos a detallar qué papel va a cumplir cada servicio externo y cómo diseñamos la interacción con estos:

5.4.1- Spotify

Esta gran plataforma digital de divulgación y reproducción musical dispone de una sección para desarrolladores que habilita el uso de ciertas de las funciones de su aplicación mediante una API de desarrollo. Esta API nos permite, desde cualquier sistema externo, dotar de funcionalidades avanzadas de tratamiento musical o incluso gestión de usuarios en Spotify.

La parte de gestión de usuario no nos interesa mucho, no obstante, al principio del estudio se consideró interesante la posibilidad de buscar playlists públicas y extraer desde ahí canciones relevantes, por ejemplo, desde listas de reproducción de música de estudio, relajación...

Al final lo que más nos ha interesado es la posibilidad de obtener los parámetros físicos de una canción, y es la característica que va a dar forma a todo este TFG. Además, la posibilidad de buscar canciones por género también es muy interesante dada la gran cantidad de géneros que tiene Spotify.

Para poder utilizar la API de desarrollador se requiere registrar un usuario en el sistema de desarrollo de Spotify, al cual se le brinda una clave de acceso que debe introducir en el proceso de autenticación de su código. De esta manera se crea un objeto de sesión que permanece hasta el final de la ejecución y habilita el uso de las funciones de Spotify. Para poder crear esta clave se requiere dar a Spotify explicaciones de por qué necesitamos este servicio y dar una estimación del uso, y ellos deben conceder el permiso o no. En nuestro caso, y probablemente en ninguno, han dado ningún problema.

Cada una de las funciones se utiliza referenciando el objeto de sesión e invocando alguna de sus funciones, es una forma muy sencilla de realizar lo que en el fondo es una llamada REST, fácil de implementar y amigable con el usuario.

Las funciones devuelven cadenas de texto en formato JSON¹⁰ que pueden entenderse fácilmente mediante el uso de cualquier librería de JSON. Esto permite hacer una sencilla extracción de parámetros y obtención de información útil.

Nosotros vamos a utilizar este servicio con dos fines, en primer lugar, servir de referencia para, a partir de los géneros establecidos, obtener una lista de canciones, recoger esa lista de canciones y, para cada una de ellas, realizar los procesos pertinentes, entre los cuales se encuentra la extracción de parámetros, de la cual también se encargará Spotify.

¹⁰ JavaScript Object Notation, es un formato de texto ligero para el intercambio de datos, se ha adoptado firmemente como alternativa a XML

5.4.2- Last.fm

La plataforma online de reproducción y categorización musical Last.fm, no demasiado conocida en nuestro país, pero ampliamente usada internacionalmente, pone a nuestra disposición una de las utilidades más referenciadas en el estudio de las emociones en la música, el “social tagging”. Como ya hemos comentado en el estudio de Youngmoo E. Kim, estas etiquetas están siendo un elemento de referencia de cara a la obtención de información social respecto a obras musicales concretas, ahorrando la necesidad de juntar a una muestra de población para someterlos a un test largo y aburrido. Como punto negativo, puede mencionarse la ausencia de seriedad que puede existir en estas muestras, no obstante, se disponen de medios para reducir o incluso eliminar las muestras inservibles, por lo que se decidió que Last.fm sería nuestra referencia.

Para poder solicitar los servicios de Last.fm es imprescindible registrar un usuario y un proyecto, para así obtener una clave que deberá enviarse junto con las llamadas para demostrar que, quién usa esas funciones, es un usuario.

Last.fm, para poder interactuar con sus servicios, dispone de una API que hace las veces de interfaz, pero para poder acceder a las funciones debemos efectuar una llamada a una URL con unos parámetros concretos (equivaldría a un método GET). Editando esa URL podemos seleccionar la función que deseemos, los parámetros que le enviamos y el formato en el que solicitamos la respuesta.

Así pues, como función le pediríamos la función toptags, que nos devuelve una lista con los tags ordenados para una canción concreta, seguido del nombre de la canción y el artista, y por último el formato que en nuestro caso sería JSON. Se prefiere este formato por la sencillez en la extracción y porque ya lo estamos utilizando para Spotify, por lo que no requerimos tratamiento especial ni importación de nuevas librerías.

Estos tags serán utilizados en nuestro sistema como base para la obtención de valores numéricos a partir de un estudio lexicográfico.

5.4.3- Indico.io

Indico¹¹ es un grupo de investigadores de Boston que han desarrollado un conjunto de funciones potentes y fáciles de usar relacionadas con el aprendizaje automático. Las funciones que brinda la API de Indico son numerosas, a la par que potentes, y tocan prácticamente cualquier ámbito, desde el reconocimiento de idiomas en un texto hasta la búsqueda de patrones en fotografías.

En nuestro caso, poseen una función que nos resulta especialmente interesante, se trata de un analizador lexicográfico que, dada una cadena de texto, devuelve valores numéricos que representan emocionalmente el estado del texto según los siguientes parámetros:

- Enfado
- Alegría
- Miedo
- Tristeza
- Sorpresa

¹¹ <https://indico.io>

El resultado será un conjunto de números, que entre todos sumarán uno, cuanto mayor sea un número más tendencia tendrá dicho texto por esa emoción concreta. Cuanto más se aleje el total de uno, menos segura será la estimación.

Esta función viene como anillo al dedo para el desarrollo del proyecto, ya que nos permite cuantificar de manera emocional los abstractos tags que, de otra manera, no nos darían demasiada información.

Como es habitual, es requerida una clave de acceso que se brinda en el registro del usuario y las especificaciones del proyecto que se lleva a cabo. En el caso de Indico, en lugar de darnos baja prioridad en las respuestas nos imponen la limitación de que solo podemos realizar 10.000 llamadas, esta condición será determinante para el desarrollo del proyecto. Dichas llamadas pueden ampliarse apoyando económicamente al grupo o solicitando un aumento por escrito.

Indico posee una API de desarrollo, pero desgraciadamente no funciona bien en nuestro sistema, por lo que para poder acceder a tan deseada función se debe descubrir y llamar manualmente al servicio mediante una POST request, sin interfaces que nos faciliten la comunicación.

En la llamada hay que especificar la clave de usuario, la función que se efectuará y los argumentos, que en este caso será la lista de tags concatenados directamente.

Con esta llamada, la intención es que compense de un solo golpe todos los tags ya seleccionados y nos devuelva los valores numéricos medios que asumiríamos, no pertenecen a un tag concreto, sino a toda la canción como producto de todos sus tags. Aunque los tags se mandan como una única cadena, Indico nos especifica que cada ítem de estudio cuenta como una llamada para esta función, es por eso que debemos ser cuidadosos con la cantidad de ítems que mandamos para evitar pasarnos del tope establecido.

La respuesta es una cadena de texto en formato JSON que sólo contiene las cinco parejas clave-valor de las emociones expuestas con su valor numérico, que nosotros almacenaremos directamente en nuestra base de datos.

5.5- Base de datos

Respecto a la base de datos, su diseño es bastante sencillo. Dispondremos de una base de datos llamada SME, que constará de una única tabla **cancion** con múltiples columnas de datos entre los cuales se guardarán campos como **nombre**, **artista**, **álbum**, **id**, los parámetros obtenidos de Spotify como **valencia**, **duración**, **tempo**, **tono**, **clave rítmica**, además de los **tags** y las **valoraciones lexicográficas de indico**. Se llega a la conclusión de que no es preciso diferenciar tablas como canción, autor, disco, las cuales, aunque son muy propensas a existir y tener relaciones entre sí, no son necesarias para nuestro proyecto.

No obstante, si nuestro sistema pretendiera acceder a las canciones mediante criterios de búsqueda entre los cuales se vean implicados procesos de asociación por autor, álbum o cualquier otro parámetro, sería necesario rediseñar la base de datos e introducir tablas adicionales que cumplan con las nuevas necesidades. Esto ralentizaría los procesos de inserción y acceso a los datos, pero agilizaría enormemente dichas búsquedas. Como no es requisito en nuestro proyecto simplemente mencionaremos esta otra posibilidad.

6- Implementación

6.1- Gestión temporal

Una vez con todos los procesos y objetivos claros, se procede a implementar el sistema planteado, en un proceso que abarca entre tres y cuatro meses de trabajo, con una asignación semanal de entre quince y veinte horas que combinan estudio de tecnologías, aprendizaje de conceptos, práctica, tests e implementación definitiva.

Siguiendo el análisis temporal y el establecimiento de unos puntos de control previamente, se destina en torno a un mes de desarrollo para cada prototipo. Un prototipo se considera terminado cuando se han realizado pruebas satisfactorias individuales para cada uno de los puntos, y, posteriormente, una prueba global de todo lo realizado hasta el momento.

Algunos de los puntos acaban retrasando nuestro desarrollo más de lo planeado debido a la inexperiencia en diversos ámbitos, y a cambios inesperados en nuestra orientación.

6.2- Decisiones de implementación

A lo largo del desarrollo del sistema, no todo fue a pedir de boca desde el principio, se dieron una serie de circunstancias que obligaron a replantear ciertos aspectos. Dichas decisiones serán listadas y explicadas a continuación:

1. En un principio no se iba a recurrir a un analizador sintáctico (Indico), sino que iban a tratarse los tags como un objeto independiente y se pasarían en masa por nuestro mecanismo de minería. No obstante, este enfoque presentaba serios problemas, en primer lugar, el volumen y la ilegibilidad de los datos introducidos, además de problemas de formato y lectura al meter directamente en la base de datos un objeto de tipo JSON parseado a texto. Este hecho obligó a suprimir este formato y a concatenar cada tag en una cadena de texto separada por comas, para su mejor legibilidad. La contraparte de este proceso fue la pérdida del atributo que muestra la repetición de cada tag. En estudios futuros puede replantearse este diseño para recuperar ese atributo y tenerlo en cuenta para diseñar modelos con peso en función de la mayor repetición de etiquetas.
2. Continuando con el problema anterior, una vez se obtuvieron las etiquetas concatenadas, lo cual permitía ver de una manera más legible el contenido de las mismas, se pudo observar que la mayor parte de las etiquetas pecan de no ser nada interesantes, ya sea por no transmitir nada nuevo o por directamente ser “basura”. Al tratarse de una fuente introducida por el usuario, y, por tanto, muy propensa a errores, se vio la necesidad de crear un filtro de etiquetas avanzado más allá de las simples limitaciones que poníamos en un principio. Dicho cambio se detalla como sigue:
 - Cada tag será estudiado individualmente
 - Determinamos que un tag que componga varias palabras será muy probablemente inútil, por lo que anularemos todo aquel que contenga un espacio. Esta decisión se toma estudiando una muestra de 100 canciones
 - También determinamos que no nos aporta nada el hecho de que un tag se refiera al género de una canción (algo mucho más frecuente de lo imaginado), por lo que se determina una lista de palabras prohibidas, y los tags que coincidan serán eliminados

- En un principio se estableció como requisito que una canción que no tuviera un mínimo de mismas tags, no podía ser candidata para nuestro estudio. Considerando que dichas etiquetas constituyen nuestra “base de verdad” y que son vitales para nuestro estudio es imprescindible que esta regla siga manteniéndose después del filtrado. De esta manera, incluso una canción que posea tags, si no son considerados como información de calidad, no será procesada.
- 3. En un punto más avanzado del desarrollo, mientras se avanzaba en las técnicas de minería que se iban a utilizar, se descubrió la librería de Indico.io, que nos brindaba funcionalidades de Big Data y de análisis lexicográfico. Se tomó la decisión de incorporarla al sistema (y se modificó por tanto el diseño para que fuera acorde al cambio). Esta incorporación trajo funcionalidades nuevas, pero determinó un nuevo diseño para la base de datos, que implicó realizar un borrado de los datos previos.
- 4. Poco después, trabajando con Indico se descubrió que, de manera gratuita, Indico solo pone a disposición de los usuarios libres un número limitado de usos, y ese número era insuficiente para cumplir con nuestros objetivos por lo que tuvimos que rediseñar nuestros límites de la siguiente forma:
 - Mil canciones en la base de datos como máximo (previamente se había acordado introducir todas las canciones posibles para tener una mayor muestra)
 - Cada canción solo puede tener hasta diez tags, por lo que a la vez que se filtra, se debe dejar de añadir tags cuando se cumplen los diez. Esto se diseñó así porque en el caso ideal Indico analizaría las diez tags en cada una de las mil canciones, cubriendo así el tope de usos que tiene un usuario cualquiera (diez mil), aunque en la realidad acaban siendo bastantes menos debido a que no todas las canciones tienen diez etiquetas válidas
- 5. En un punto avanzado del desarrollo se descubrió que la API de Spotify había añadido una función que al principio del estudio no estaba, la cual nos permitía elegir un género y buscar canciones relevantes asociadas a ese género. Este hecho supuso un replanteamiento radical, ya que previamente se estaba utilizando la función de “Temas actuales”, que devolvía un reducido espectro de estilos diferentes, y además suelen ser canciones de “usar y tirar”, lo cual hace que no sean poseedoras de tags en la mayor parte de casos, al ser temas nuevos que viven poco tiempo. Esta función nueva nos permite determinar un vector de estilos y pasarlos todos a la vez, por lo que, para ajustarnos al plan limitado que Indico nos presenta, elegiremos diez estilos musicales diferentes y obtendremos cien canciones relevantes distintas por estilo, cubriendo así las mil canciones estipuladas.

- Los estilos seleccionados para la muestra han sido elegidos por el estudiante, se han buscado estilos tanto actuales como históricos, diferentes entre ellos e intentando que abarquen al máximo de población posible, de tal manera que las muestras de datos sean capaces de demostrarnos que, entre estilos, también hay diferencias. Los estilos son los siguientes:
 - i. Rock
 - ii. Pop
 - iii. Funk
 - iv. Blues
 - v. Electronic
 - vi. Metal
 - vii. Sad
 - viii. Latino
 - ix. Grunge
 - x. Soul
- Dichos estilos son nominados de esta manera en Spotify, no se refieren exclusivamente a estilos musicales como podemos apreciar en el estilo “sad”, no obstante, consideraba interesante introducir un estilo que Spotify ya tiene directamente asociado a un estado emocional. De la misma manera existen estilos como relaxing, gaming, studying y música situacional para diferentes momentos
- Aunque este tipo de categorizaciones van muy en la línea de nuestra investigación y serían de gran apoyo en nuestro proyecto, el hecho de que utilizemos etiquetas de Last.fm para hacer el análisis nos priva de la posibilidad de usar toda esta música ambiental, que en enorme medida no es escuchada ni etiquetada al no ser, generalmente, una música demasiado popular. Es por eso que la música que más nos ayuda tiene la forma de canción típica, que un usuario cualquiera buscaría en Last.fm y etiquetaría. Esta decisión parte de la experiencia de haber analizado más de cien canciones diferentes en varios estilos emocionales y haber conseguido que solo un par de ellas pasen el filtro, canciones que suelen ser las más populares de toda la muestra.
- No se ha escogido el estilo “clásico” debido a que todos los estudios que han trabajado de manera parecida han utilizado generalmente música clásica, además de que es un estilo con una enorme variedad emocional y muy dispar en sus especificaciones, así pues, una obra barroca no tiene nada que ver con una obra impresionista, ni compositivamente ni emocionalmente, sin embargo, las dos están categorizadas dentro del estereotipo de “música clásica”. Por esta razón se ha decidido omitir ese estilo y utilizar otros más rígidos y más contemporáneos que tratan de obtener los mismos resultados, siendo consciente por supuesto de que en muchos estilos nos encontramos con posturas emocionales diferentes según el tema musical (pudiendo encontrar por ejemplo temas de Rock triste, Rock enérgico, Rock melancólico...)

Una vez ya se sortearon todas las dificultades, se completó la implementación y, tras los test pertinentes, se comprobó que era completamente funcional, por lo que se procedió a su ejecución.

6.3- La ejecución determinante

Con todos los aspectos individuales funcionando correctamente y un algoritmo funcional, había que prepararlo para el proceso de inserción en masa al que íbamos a someter al sistema.

Mil canciones pueden no ser muchas, pero cuando para cada canción actúan varios procesos que implican llamadas REST remotas, análisis de cada etiqueta por separado, acceso a una base de datos y el “commit”, el tiempo de ejecución empieza a crecer de una forma desmedida, principalmente por la cantidad de llamadas externas que deben realizarse de modo secuencial ya que se necesitan los resultados de cada una para poder acceder a la siguiente.

Este hecho, sumado a que disponemos de un número limitado de llamadas a la API de Indico, nos obligan a realizar el proceso de inserción sin desperdiciar una sola llamada, y el hecho de que gaste tanto tiempo favorece prepararlo para una sola ejecución que dure las horas que necesite.

Para ello era importante que el sistema:

- No se detuviese ante los errores de inserción
- No analice las canciones que ya hayan podido ser analizadas (por eso existe el diccionario de claves)
- No analice los tags que no van a insertarse
- Muestre siempre el progreso en la consola

El hecho de que no se haya diseñado en las especificaciones, un mecanismo de persistencia que nos permita conservar el diccionario de claves entre ejecuciones del sistema nos obliga a realizar el proceso de golpe sobre una base de datos vacía (si alguna canción ya se ha introducido en ejecuciones anteriores no tenemos manera de identificarla más que buscándola en la base de datos, lo cual añade un nivel de complejidad más, y además demora más tiempo en las últimas iteraciones). Si lo hiciéramos en varias ejecuciones diferentes correríamos el riesgo de analizar muchas veces temas repetidos y agotar estúpidamente los recursos disponibles.

Para conseguir el éxito en esta operación, era importantísimo que todos los puntos de los prototipos se hubieran probado minuciosamente, de tal manera que el sistema no abortara la ejecución y fuera capaz de gestionar todas las entradas de forma eficaz. Para poder llegar a este punto hubo que enfrentarse a varios problemas.

En primer lugar, debido a que no distinguimos canciones por país (cosa que limitaría muchísimo la completitud del estudio), los títulos de las canciones contendrán caracteres especiales para nuestra codificación, lo cual obligó a adoptar una codificación capaz de representar cualquier símbolo Unicode (UTF-8), y a llevar a cabo labores de prevención y de escape (utilización de un carácter que modifica el funcionamiento del carácter subsecuente) de caracteres especiales, que chocaban con la codificación y generaban errores sintácticos.

El sistema perdía demasiado tiempo realizando todas las verificaciones en algunos casos donde ya se había determinado que un objetivo no era válido. Este problema tuvo fácil solución, ya que con un poco más de experiencia por parte del desarrollador se pudieron diseñar algoritmos más eficientes que realizaban estas funciones de manera óptima, reduciendo el tiempo de ejecución potencialmente en varias horas.

Había que enfrentarse a una decisión de equilibrio: El sistema abre una conexión con la base de datos al comienzo de la ejecución y para guardar los cambios en ésta debe realizarse un commit. Este commit gasta un tiempo en ejecución y lo idóneo sería realizarlo al terminar el algoritmo, pero si el sistema se aborta o se detiene de manera inesperada durante la ejecución, todos los resultados se perderían, por lo que se pensó en ir realizando commits cada vez que se introducía un dato, aunque ralentizara el resultado final, ya que es mucho más seguro que hacerlo al final. Un corte en la luz o un error repentino no gestionado podía poner fin a la inserción y reducir drásticamente nuestra muestra de datos disponible.

Con todo esto en cuenta, se llevó a cabo la ejecución en el entorno más estable y a prueba de errores posible.

- En un ordenador portátil con batería para prevenir cortes eléctricos
- Con secuencias try-catch en cada sección crítica
- Probando previamente que cada uno de los servicios web requeridos se encuentra disponible

Esta ejecución tuvo éxito casi absoluto, y como resultado final, ahora disponemos de una base de datos con novecientos ochenta elementos de tipo canción, todos ellos con una lista de tags suficiente para estudiar y con los parámetros físicos de la canción, además de la valoración lexicográfica de los tags.

7- Minería de datos

Con todo el sistema funcionando y una base de datos llena de elementos interesantes para el TFG, el objetivo de esta sección es realizar un pequeño estudio que refleje la utilidad o no de dichos datos.

Todo el desarrollo de este proyecto se ha efectuado sobre una base o “suelo de verdad” sintético, el propio concepto del proyecto es abstracto, y debido a que el estado del arte todavía no es capaz de generar modelos precisos [Schuller, Dorfner, Rigoll 2010], damos por hecho que este estudio no será capaz de arrojar resultados prometedores y definitivos.

Teniendo claro que la base de este estudio podría no ser la adecuada, se vuelve interesante que para cerrar este informe trabajemos en procesos sencillos de minería y análisis estadístico con las muestras de datos de que disponemos. Estos estudios podrían acercarnos un poco más a la utilidad real de lo aquí desarrollado y la implicación que tendría en el avance de la línea de investigación.

Al final, se trata de ver si las fuentes elegidas y la forma de combinar las propiedades musicales tiene sentido y nos puede mostrar patrones de comportamiento, correlaciones, atributos significativos, parámetros insignificantes... dentro de nuestras muestras, y demostrar que los planteamientos propuestos por [Kim, Schmidt, Migneco, Richardson, Scott, Speck y Turnbull 2010] son válidos y aportan algo a la investigación.

Debido a que no se poseen los medios suficientes para realizar un gran proceso de minería y clasificación, se elegirá un pequeño entorno de desarrollo particular donde realizaremos pruebas y clasificaciones básicas.

El lenguaje elegido para construir el sistema de minería es el lenguaje R, este lenguaje enfocado en el análisis estadístico es potente, sencillo de usar y muy cómodo de cara a utilizar sus librerías y funciones. Junto con su IDE RStudio, ponen a nuestra disposición un conjunto de herramientas que nos servirán para realizar este apartado.

En primer lugar, se requiere conectar el sistema de minería con la base de datos MySQL donde guardamos toda la información. Una vez conectada, realizaremos una query que nos permita obtener toda la muestra de datos y guardarla como un tipo dataset que R pueda entender y trabajar.

Una vez con esa muestra ya se trata de aplicarle procesos y funciones predefinidos de las librerías de R. Una de las funciones más interesantes que podemos aplicar es la de resumen, con la cual el sistema nos hace un análisis de todos los atributos de la base y nos los muestra ordenadamente, con sus medias, sus cuartiles, máximos y mínimos, donde sea aplicable. El resultado puede observarse en la siguiente figura:

Results:													
name	artist	album	duration	id	energy	liveness	tempo	speechiness	acousticness	timesig	danceability	tone	loudness
Length:500	Length:500	Length:500	Min. : 40133	Length:500	Min. :0.0581	Min. :0.01880	Min. : 63.09	Min. :0.02420	Min. :0.0000021	Min. :1.000	Min. :0.1430	Min. : 0.00	Min. : -22.859
Class :character	Class :character	Class :character	1st Qu.:207896	Class :character	1st Qu.:0.5290	1st Qu.:0.08887	1st Qu.:101.14	1st Qu.:0.03360	1st Qu.:0.0065475	1st Qu.:4.000	1st Qu.:0.4520	1st Qu.: 2.00	1st Qu.: -9.598
Mode :character	Mode :character	Mode :character	Median :242740	Mode :character	Median :0.7200	Median :0.12200	Median :121.43	Median :0.04495	Median :0.0792000	Median :4.000	Median :0.5585	Median : 5.00	Median : -6.656
			Mean :249579		Mean :0.6752	Mean :0.17780	Mean :122.76	Mean :0.06498	Mean :0.2135522	Mean :3.896	Mean :0.5610	Mean : 5.38	Mean : -7.323
			3rd Qu.:276063		3rd Qu.:0.8465	3rd Qu.:0.22500	3rd Qu.:140.05	3rd Qu.:0.06982	3rd Qu.:0.3440000	3rd Qu.:4.000	3rd Qu.:0.6783	3rd Qu.: 9.00	3rd Qu.: -4.862
			Max. :544627		Max. :0.9970	Max. :0.96400	Max. :207.57	Max. :0.41100	Max. :0.9880000	Max. :5.000	Max. :0.9560	Max. :11.00	Max. : -1.665
mode	style	tags	anger	joy	fear	sadness	surprise						
Min. :0.00	Length:500	Length:500	Min. :0.0008918	Min. :0.0127	Min. :0.004135	Min. :0.004135	Min. :0.005394						
1st Qu.:0.00	Class :character	Class :character	1st Qu.:0.0116213	1st Qu.:0.5187	1st Qu.:0.024832	1st Qu.:0.024832	1st Qu.:0.028591						
Median :1.00	Mode :character	Mode :character	Median :0.0258380	Median :0.7281	Median :0.043191	Median :0.043191	Median :0.054291						
Mean :0.63			Mean :0.0473199	Mean :0.6674	Mean :0.063368	Mean :0.063368	Mean :0.083467						
3rd Qu.:1.00			3rd Qu.:0.0510414	3rd Qu.:0.8514	3rd Qu.:0.080153	3rd Qu.:0.080153	3rd Qu.:0.104246						
Max. :1.00			Max. :0.4927440	Max. :0.9584	Max. :0.323240	Max. :0.323240	Max. :0.485131						

Ilustración 6:Análisis preliminar de los parámetros a estudiar

Según estos resultados ya podemos confirmar varias cosas al respecto de nuestra muestra de datos. Algunos de estos atributos no tienen sentido en un estudio como éste, como los campos alfanuméricos, la clave rítmica, el tono o el modo, pero todos los demás pueden ser interesantes.

Para empezar, visualizando las cinco categorías de Indico (anger, joy, fear, sadness y surprise), observamos que en general prima la alegría, teniendo una media alta y una máxima casi absoluta, y que otras como miedo o tristeza no suelen recibir altas valoraciones. Esto es un hecho curioso dado que entre nuestra muestra se encuentran canciones seleccionadas directamente del estilo “sad” de Spotify, no obstante, las etiquetas de Last.fm y el posterior análisis de esas etiquetas parecen no compartir este enfoque.

Los parámetros físicos no nos dicen mucho, y deberemos cruzarlos con otros parámetros para intentar obtener correlaciones. No obstante, realizando cruces básicos y asociaciones entre pares de parámetros tampoco conseguimos ver correlaciones directas muy fuertes. Habría que estudiar más a fondo la muestra con mecánicas más potentes entre grupos de parámetros para sacar información relevante.

Como punto fuerte del estudio, se realizarán diferentes pruebas de clusterización K-means agrupando diferentes parámetros para buscar correlaciones entre los parámetros que más probablemente se relacionen.

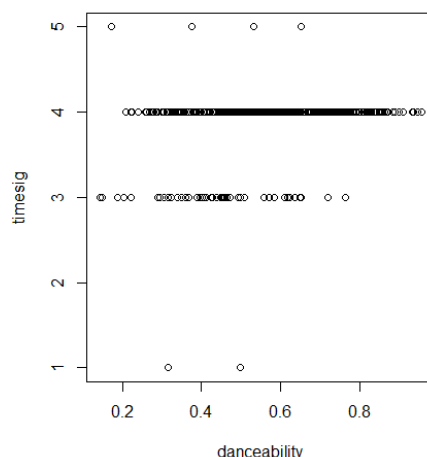


Ilustración 7: Clusterización que agrupa bailabilidad con medida rítmica (timesig, abreviación de “time signature”)

En la ilustración 7 observamos cómo en general, no podemos establecer reglas con los tiempos de 4 pulsos, no obstante, en general, los ritmos de 3 y 5 pulsos se encuentran por debajo de la media. Esto se debe sobre todo a la proliferación de las medidas de cuatro pulsos en la música actual e histórica, y normalmente ritmos diferentes suelen pertenecer a temas musicales exóticos o progresivos. En todo caso, se plantea como trabajo futuro interesante buscar temas folclóricos o clásicos con estilo “valse” y comprobar la valoración que obtienen en bailabilidad, o estudiar canciones con ritmo de zortziko¹².

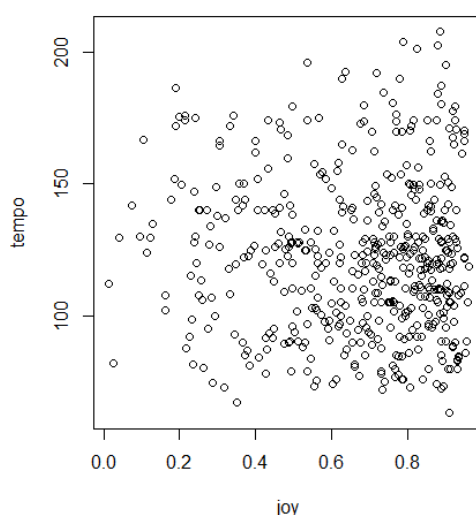


Ilustración 8: Clusterización entre alegría y tempo

Como podemos observar en la ilustración 8, la muestra de datos es heterogénea y no es sencillo encontrar correlaciones directamente. No obstante, podemos apreciar una tendencia, que los

¹² Ritmo vasco tradicional con compás de 5/8, que sería de los pocos catalogados con un 5 en “time signature”

mayores valores de alegría se encuentran, sobre todo, entre los tempos de 110 y 140 BPM (Beat Per Minute).

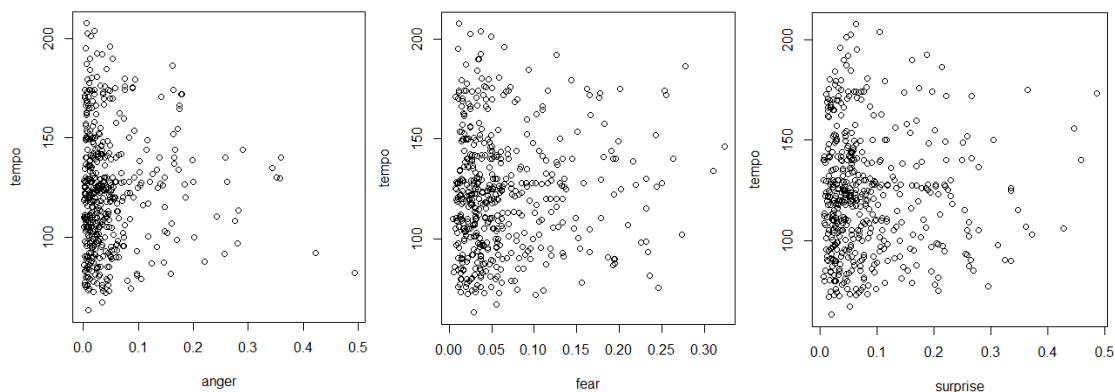


Ilustración 9: Comparación entre los resultados de enfado, miedo y sorpresa respectivamente

No obstante, en la ilustración 9 podemos observar que generalmente, nuestro analizador lexicográfico no asigna grandes valores al resto de parámetros. Este hecho puede llevarnos a pensar que las etiquetas más comunes suelen caer en la repetitividad de expresar emociones o sensaciones poco precisas. Un buen ejemplo de esto sería coger una canción “muy triste” y otra “muy alegre”, probablemente en los dos casos, un usuario al que le guste dicha canción podría escribir la etiqueta “amazing”, la cual tiene un componente medio-bajo de sorpresa y un gran componente de alegría según el analizador. En este supuesto, no estaríamos captando la totalidad del carácter triste de la canción a partir de los tags.

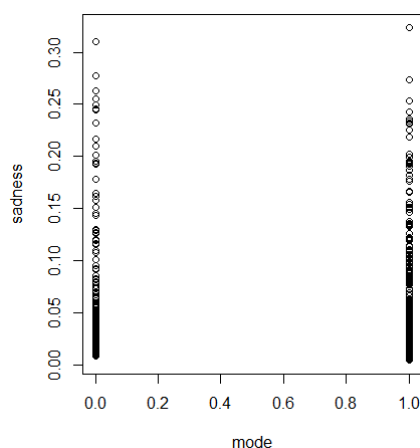


Ilustración 10: Valores de tristeza según el modo musical

Una de las cuestiones tradicionales en música es la asociación de la escala menor con formas más oscuras, más tristes, menos joviales. En este estudio se ha intentado realizar esta comparación como se observa en la ilustración 10. Sin embargo, el resultado dista mucho de ser satisfactorio, al encontrar los valores más altos precisamente en el modo mayor (valor de 1 en modo).

La razón detrás de esto es probablemente la misma que perjudicaba la anterior comparación, al haber muy pocas referencias con un auténtico valor de tristeza. Estos dos casos son

preocupantes ya que en la muestra hay 100 canciones seleccionadas directamente del género “sad” de Spotify.

Para solventar este problema una posible solución sería adoptar uno de los enfoques de Meyers y añadir estudio de líricos a nuestro sistema. El planteamiento sería muy interesante y podría ayudar a verificar la integridad del atributo “speechiness”, que nunca alcanza grandes números, o fundamentalmente, con un estudio lexicográfico como el que nos puede proporcionar Indico, obtener una segunda valoración de los parámetros emocionales del texto, pero esta vez estudiando la letra de una canción.

Por lo demás, lamentablemente este tipo de asociaciones nos son insuficientes para encontrar relaciones interesantes en la muestra de datos. La subjetividad de las etiquetas, así como la poca precisión de las mismas, nos dejan unos resultados vagos que no terminan de captar muy bien las características de las canciones. Con unos resultados tan dispersos, resulta inviable plantear modelos de predicción.

Con procesos de minado más exhaustivos, redes neuronales y aplicación de mecanismos de aprendizaje automático probablemente estas mismas muestras de datos puedan reflejar resultados diferentes y más esperanzadores, y tal vez, mejorar la precisión de en torno al 60% que estaban consiguiendo los últimos sistemas.

8- Licencia software y documental

Copyright (c) 2017, Guillermo Ramos Sanz

Todos los derechos reservados.

La redistribución y el uso en las formas de código fuente y binario, con o sin modificaciones, están permitidos siempre que se cumplan las siguientes condiciones:

1. Las redistribuciones del código fuente deben conservar el aviso de copyright anterior, esta lista de condiciones y el siguiente descargo de responsabilidad.
2. Las redistribuciones en formato binario deben reproducir el aviso de copyright anterior, esta lista de condiciones y el siguiente descargo de responsabilidad en la documentación y/u otros materiales suministrados con la distribución.

ESTE SOFTWARE SE SUMINISTRA POR GUILLERMO RAMOS SANZ “COMO ESTÁ” Y CUALQUIER GARANTÍAS EXPRESA O IMPLÍCITA, INCLUYENDO, PERO NO LIMITADO A, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN Y APTITUD PARA UN PROPÓSITO PARTICULAR SON RECHAZADAS. EN NINGÚN CASO GUILLERMO RAMOS SANZ O COLABORADORES SERÁN RESPONSABLES POR NINGÚN DAÑO DIRECTO, INDIRECTO, INCIDENTAL, ESPECIAL, EJEMPLAR O COSECUENCIAL (INCLUYENDO, PERO NO LIMITADO A, LA ADQUISICIÓN O SUSTITUCIÓN DE BIENES O SERVICIOS; LA PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS; O INTERRUPCIÓN DE LA ACTIVIDAD EMPRESARIAL) O POR CUALQUIER TEORÍA DE RESPONSABILIDAD, YA SEA POR CONTRATO, RESPONSABILIDAD ESTRICTA O AGRAVIO (INCLUYENDO NEGLIGENCIA O CUALQUIER OTRA CAUSA) QUE SURJA DE CUALQUIER MANERA DEL USO DE ESTE SOFTWARE, INCLUSO SI SE HA ADVERTIDO DE LA POSIBILIDAD DE TALES DAÑOS.

Las opiniones y conclusiones contenidas en el software y la documentación son las de los autores y no deben interpretarse como la representación de las políticas oficiales, ya sean expresas o implícitas, de Guillermo Ramos Sanz.



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

9- Conclusiones y Trabajo Futuro

Este TFG se realizó en un principio intentando dar soporte y formar parte de un proyecto nacional destinado a investigar esta línea y continuar con este desarrollo en un entorno de trabajo profesional organizado y con una meta ya centrada en el usuario. La verdad es que este sistema podría ser en multitud de sistemas de apoyo a la toma de decisiones.

Sin ir más lejos, el objetivo inicial del proyecto nacional que se aprobó en Zaragoza, era la de diseñar un sistema de recomendaciones que, según una serie de sensores capaces de captar la emoción de un deportista, fuera capaz de estimular positivamente al deportista eligiendo la canción adecuada. Este planteamiento en el mejor de los casos supondría un aumento en el rendimiento de dicho deportista.

Pero no sólo sería viable en ese caso, existen multitud de entornos donde podría ser interesante un sistema así. Por ejemplo, se podría pensar en un sistema de selección automática de música para gimnasios, en función del plan de entrenamiento propuesto, o bien un sistema de apoyo a las personas que han sufrido algún tipo de accidente o problema emocional, pudiendo servir para el desarrollo de sesiones de musicoterapia. Las posibilidades son muy grandes y todo pasa por la capacidad de extraer valor dentro de las muestras seleccionadas, y más importante, de entender que realizando estas asociaciones de datos puede obtenerse información útil para el usuario final, tal vez no directamente, pero sí formando parte de un sistema más grande.

Es cierto que los resultados obtenidos por sí solos no son demasiado llamativos debido a que la minería realizada no ha sido muy profunda, y tampoco se han podido asociar todos los parámetros entre sí directamente. También es cierto que hasta ahora no se han conseguido sistemas de predicción en este campo que fueran más precisos que en torno a un 60%, y probablemente este sistema tampoco lo sea, no obstante, eso no quiere decir que este modelo sea inviable, sino que requiere una mayor profundidad de análisis.

En cuanto a trabajo futuro, existen muchas posibilidades para este TFG, entre las cuales figura el formar parte del proyecto nacional mencionado anteriormente, así como cualquier otra línea de investigación relacionada, que obligarían a desarrollar mejoras para este sistema, entre las cuales podrían fácilmente encontrarse las siguientes:

- Mejorar el sistema de introducción, recopilar más fuentes, prevenir fuertes dependencias de una sola tecnología, buscar sistemas similares
- Actualizar el sistema de bases de datos. Ahora mismo es simplemente un diseño de prueba, un despliegue real para una aplicación real distaría mucho de esta propuesta teórica. Con ese fin sería necesario contar con una base de datos replicada, con diferentes protocolos de acceso, seguridad y copias de seguridad. Además, el diseño es muy básico por lo que habría que ampliarlo y considerar el migrar a una base de datos NoSQL
- Aunque se han diseñado múltiples modos para explotar Spotify e insertar canciones de varias maneras, el sistema podría mejorarse para permitir también que un usuario cargue canciones directamente al sistema, ya sea desde playlists, desde búsquedas directas... este enfoque estaría más orientado al uso directo por un usuario final
- Profundizar en los procesos de minería de datos. Sería una labor importante, y requerida en el caso de buscarle la precisión al sistema, la posibilidad de buscar correlaciones y poder definir reglas de búsqueda sería conveniente para cualquier sistema que pudiera hacer uso de esta tecnología

- Como broche final, el sistema debería migrarse a una plataforma de computación distribuida para poder beneficiarse de los privilegios de la computación paralela, mejorar los tiempos de respuesta, nuevas políticas de seguridad, eficiencia...

Como conclusión, aunque el avance que se ha conseguido realizar no sea muy vistoso, debido sobre todo a la asignación temporal de un TFG, así como a los procesos de aprendizaje y al haber tenido que lidiar con la implementación del sistema, el proyecto me ha instruido en múltiples disciplinas, ha profundizado en un aspecto actual que forma parte de una línea de investigación en desarrollo y fundamentalmente, me ha permitido trabajar en una disciplina que me gusta y me interesa. Como resultado me siento satisfecho con el trabajo realizado. Considero que el trabajo de análisis y recogida de información fue fructífero, se identificó un punto en el que no se había trabajado y se plantearon los mecanismos para trabajar ese punto. Incluso se diseñó un prototipo de sistema de ayuda a la toma de decisiones, y se analizó la viabilidad del modelo seleccionado para este problema. Para mí la experiencia ha sido grata y me ha permitido aprender y complementar mi formación con las múltiples disciplinas diferentes que se han visto implicadas en el desarrollo de este TFG.

10- Referencias Bibliográficas

Gabrielsson, A. Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, 2002.

Gabrielsson, A. y Lindström, E. The influence of musical structure on emotional expression. In *Music and Emotion: Theory and Research*, P. N. Juslin and J. A. Sloboda Eds., Oxford University Press, Oxford, UK. 2001.

Hevner, K. Experimental studies of the elements of expression in music. *American Journal of Psychology*, 1936.

Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, J. A. y Turnbull, D. Music Emotion Recognition: A State of the art review. 11th International Society for Music Information retrieval Conference (ISMIR) 2010

O. C. Meyers, A mood-based music classification and exploration system, TFM, Massachusetts Institute of Technology. 2007.

Schuller, B., Dorfner, J., y Rigoll, G. Determination of nonprototypical valence and arousal in popular music: Features and performances. *EURASIP J. Audio, Speech, Music Process.* 2010.

Weld, H. P. An experimental study of musical enjoyment. *American Journal of Psychology*, 1912.

Wedin, L. A multidimensional study of perceptual– emotional qualities in music. *Scandinavian Journal of Psychology*, 1972.

Yang, Y. y Chen, H. Machine Recognition of Music Emotion: A Review. Article in *ACM Transactions on Intelligent Systems and Technology*. ResearchGate. 2012.

Zentner, M., Grandjean, D. y Scherer, K. R. Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *American Psychological Association*, 2008.

<http://www.serfelices.org/index.php/psicologia/410-el-poder-de-la-musica-sobre-las-emociones>

<http://www.rtve.es/television/20111009/musica-emociones-neurociencia/465379.shtml>

<http://www.bbc.co.uk/rd/projects/making-musical-mood-metadata>

<http://www.rdatamining.com/>

<https://www.python.org/>

<https://www.spotify.com/es/>

<https://spotipy.readthedocs.io/en/latest/>

<https://www.last.fm/api>

<https://indico.io/docs>

<http://guidetodatamining.com/>

http://www.cs.wustl.edu/~jain/cse570-15/ftp/iot_ml/

<http://music.ece.drexel.edu/research/emotion>