



Music4all + Extrasensory Recomendador musical contextualizado (Data Wrangling)

QUICAHÑO MIRANDA, Victor Alejandro

Docente: Prof Mag. VALDIVIA CUADROS, Ana María
Curso: Tópicos en Ciencia de Datos

UNSA - Universidad Nacional de San Agustín de Arequipa
Junio de 2025

Abreviaturas

Índice

1. Contexto	6
2. Análisis del comportamiento de los datos	8
2.1. Music4All [1]	8
2.2. Music4All - Onion [2]	10
2.3. Extrasensory [3]	11
2.4. Data Faltante	14
2.4.1. Ausencia de datos en Extrasensory	14
2.5. Granularidad Temporal	14

Índice de cuadros

1.	Descripción de atributos del conjunto de datos musicales	17
2.	Descripción de archivos de características musicales	18
3.	Descripción de los archivos y atributos utilizados en el análisis	19
4.	Archivos relacionados con metadatos generados por usuarios	19
5.	Archivos con representaciones visuales extraídas de videos musicales	20
6.	Resumen de archivos de historial de escucha	20
7.	Descripción de modalidades, cantidad de variables y características del dataset	21

Índice de figuras

1.	Pipeline de la elaboración de Music4All.	6
2.	Pipeline de la elaboración de Music4All.	7
3.	Capas de Music4All-Onion.	8
4.	Porcentaje de datos completos entre todos los uaurios de extrasensory. . . .	15
5.	Comparativa temporal con la energía de las canciones escuchadas en 2020 por este usuario.	16

1. Contexto

La base de datos utilizada en este proyecto está compuesta por la integración de tres datasets principales: *Music4All* [1], *Music4All-Onion* [2] y *Extrasensory* [4]. Como se muestra en la Figura 1, los datasets *Music4All* y su extensión *Onion* se complementan al incorporar características multimodales adicionales sobre las mismas pistas musicales. Por otro lado, el dataset *Extrasensory*, que proporciona información contextual y sensorial de usuarios, puede integrarse con los registros de escucha de *Music4All* mediante una unión basada en las marcas temporales (*timestamp*), lo que permite contextualizar las interacciones musicales con datos del entorno del usuario.

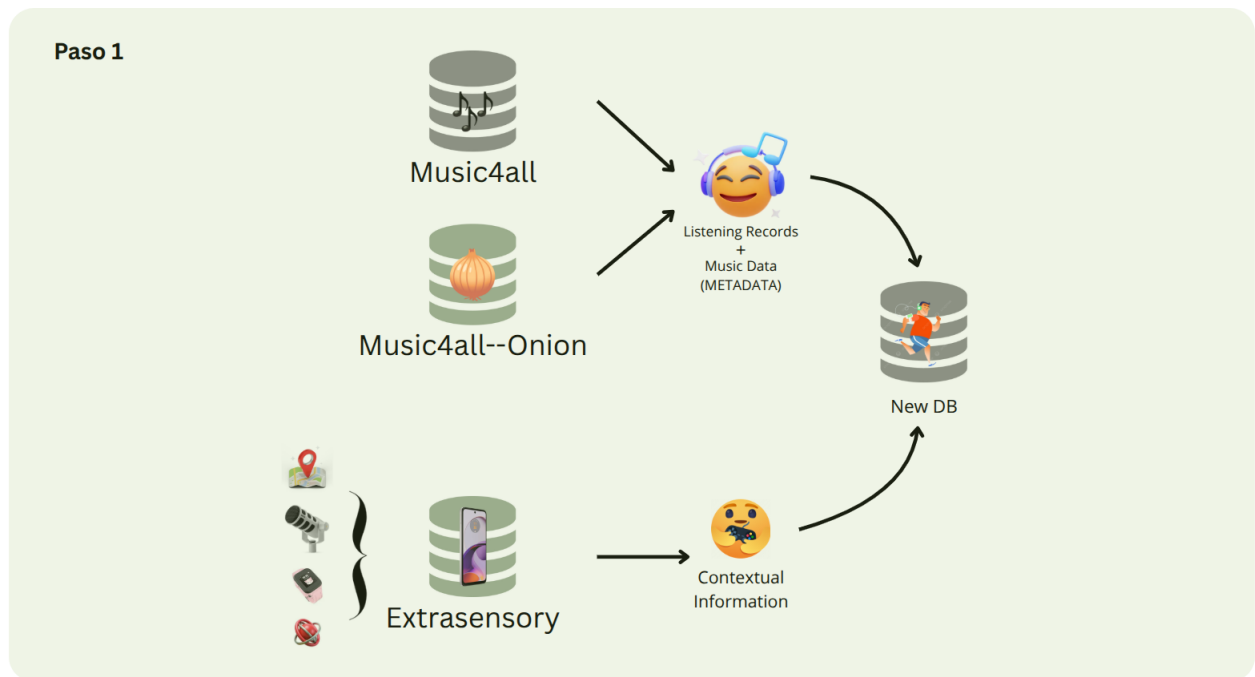


Figura 1: Pipeline de la elaboración de Music4All.

Music4all Este dataset surge como una contribución a la comunidad de Music Information Retrieval (MIR), cuyo objetivo principal es el desarrollo de sistemas que permitan recuperar y recomendar contenido musical de manera eficiente y efectiva. Pese a los avances en el área, existía una carencia de bases de datos que cumplieran con ciertos requisitos fundamentales para la investigación:

- Disponibilidad de señales de audio
- Diversidad de atributos musicales
- Un volumen considerable de piezas musicales.

Este cual ofrece un conjunto de datos amplio, más de **109,000 canciones** y el historial de escucha de **15,602 usuarios anónimos**. El dataset contiene diversos tipos de

información para cada canción, incluyendo metadatos, clips de audio de 30 segundos, letras, información de popularidad y atributos acústicos derivados de la API de **Spotify**, así como etiquetas de género y tags generados por usuarios en **Last.fm**.

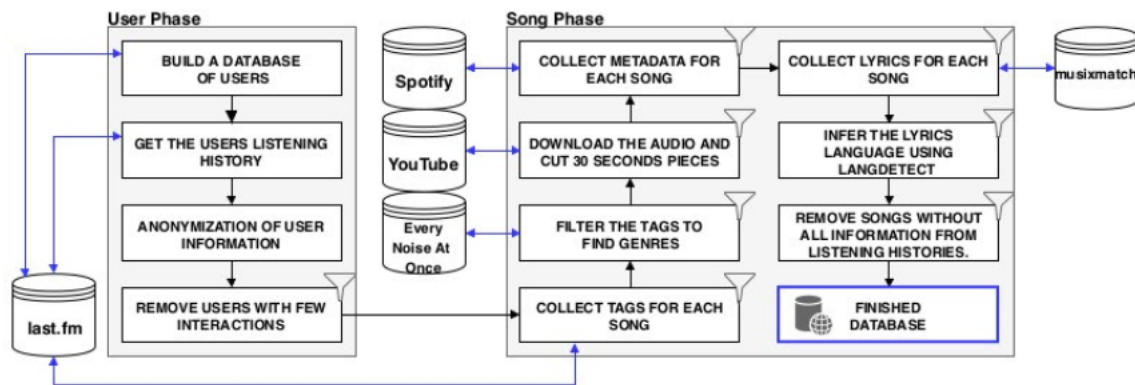


Figura 2: Pipeline de la elaboración de Music4All.

Music4All-Onion Es una extensión multimodal y a gran escala del dataset Music4All. Este fue diseñado para suplir las limitaciones comunes en los sistemas de recomendación musical actuales: la dependencia exclusiva de interacciones usuario-item (colaborative filtering), ignorando el contenido musical como base para la afinidad del usuario. Esta omisión se debe principalmente a la falta de datasets estandarizados que integren tanto información colaborativa como de contenido.

Para solucionar este problema, Music4All-Onion incorpora 26 nuevas características de tipo audio, video y metadatos, además de proporcionar 252,984,396 registros de escucha pertenecientes a 119,140 usuarios, extraídos de Last.fm. Esto permite la investigación comparativa del impacto de distintas capas semánticas del contenido musical (modelo tipo cebolla, como se muestra en la Figura 3) sobre la precisión, novedad y equidad de los sistemas de recomendación.

Extrasensory El dataset ExtraSensory fue recopilado “en la vida real” mediante la aplicación móvil ExtraSensoryApp desarrollada en la UC San Diego. Durante aproximadamente un año (2015–2016), 60 voluntarios llevaron su smartphone y en algunos casos, también un smartwatch configurados para registrar automáticamente muestras de datos sensoriales cada minuto. Al mismo tiempo, los usuarios auto-reportaban su actividad y su contexto (por ejemplo, “caminando”, “sentado”, “en reunión”, “en casa”) mediante una interfaz de la app. Este diseño permite estudiar el reconocimiento de contexto humano y la actividad diaria bajo condiciones naturales, superando las limitaciones de los experimentos controlados en laboratorio.

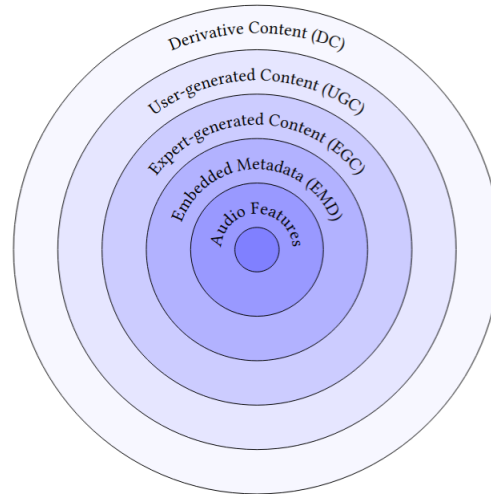


Figura 3: Capas de Music4All-Onion.

2. Análisis del comportamiento de los datos

2.1. Music4All [1]

La entidad principal de estudio en Music4All es la **canción**. Cada canción está descrita por una serie de atributos que pueden agruparse en categorías como metadatos, características acústicas, y etiquetas semánticas [**Revisar la estructura del dataset en el Cuadro: 1**]

- id
 - **Tipo de dato:** text (cadena de caracteres).
 - **Descripción:** Identificador único de 16 caracteres asignado a cada canción dentro de la base de datos. Permite referenciar de manera inmutable y sin colisiones cada registro de canción.
- artist
 - **Tipo de dato:** text.
 - **Descripción:** Nombre del artista tal como aparece en Last.fm. Hay un total de 16 269 artistas únicos. Se utiliza para agrupar o filtrar canciones por intérprete o agrupación musical.
- song
 - **Tipo de dato:** text.
 - **Descripción:** Título de la canción. No necesariamente es único (puede haber varias canciones con el mismo nombre pero diferente artista).
- lang
 - **Tipo de dato:** text.
 - **Rango / Valores posibles:** 46 idiomas distintos.
 - **Descripción:** Idioma detectado de la letra de la canción, obtenido mediante la herramienta `langdetect`. Indica en qué idioma (o principal) se encuentran las líricas.

`spotifyid` ■ **Tipo de dato:** text.

- **Descripción:** Identificador que utiliza la plataforma Spotify para referirse a esta canción. Útil para recuperar metadatos adicionales o para enlazar con la API de Spotify.

`popularity` ■ **Tipo de dato:** int.

- **Rango:** 0 a 100.
- **Descripción:** Medida de popularidad calculada por Spotify en función del número de reproducciones recientes y su rapidez. Un valor más alto indica que la canción es más reproducida y/o más actual.

`album name` ■ **Tipo de dato:** text.

- **Descripción:** Nombre del álbum al que pertenece la canción. Hay 38 363 álbumes distintos. Permite agrupar canciones por trabajo discográfico.

`release` ■ **Tipo de dato:** int.

- **Rango:** Año de lanzamiento (por ejemplo, 1960–2020).
- **Descripción:** Año en que se publicó la canción originalmente. Útil para análisis temporales de tendencias musicales.

`danceability` ■ **Tipo de dato:** double.

- **Rango:** 0.0 a 1.0.
- **Descripción:** Medida proporcionada por la API de Spotify que indica cuán adecuada es la canción para bailar. Se calcula mediante una combinación de elementos rítmicos, tempo y regularidad de la misma.

`energy` ■ **Tipo de dato:** double.

- **Rango:** 0.0 a 1.0.
- **Descripción:** Indicador perceptual de la intensidad y actividad de la canción. Valores altos suelen corresponder a sonidos más “enérgicos” o ruidosos.

`key` ■ **Tipo de dato:** int.

- **Rango:** 0 (C) a 11 (B) según notación Pitch Class.
- **Descripción:** Clave tonal principal de la canción (Tonic), expresada en la notación estándar de *Pitch Class*. Permite estudiar características armónicas.

`mode` ■ **Tipo de dato:** boolean (0 o 1).

- **Descripción:** Modalidad de la canción: 1 = modo mayor, 0 = modo menor. Indica si la pieza musical está en escala mayor o menor.

`valence` ■ **Tipo de dato:** double.

- **Rango:** 0.0 a 1.0.
- **Descripción:** Medida de positividad emocional proporcionada por Spotify. Valores cercanos a 1.0 indican música más alegre o positiva; valores cercanos a 0.0 indican música más triste o sombría.

- tempo ▪ **Tipo de dato:** double.
- **Unidad:** Beats per Minute (BPM).
- **Descripción:** Velocidad o ritmo de la canción medida en BPM. Fundamental para analizar la intensidad rítmica y la energía percibida.
- genres ▪ **Tipo de dato:** list<text>.
- **Descripción:** Lista de etiquetas de género asociadas a la canción (por ejemplo, *rock*, *pop*, *jazz*, etc.). Hay un total de 853 géneros únicos en la base.
- tags ▪ **Tipo de dato:** list<text>.
- **Descripción:** Etiquetas (tags) generadas y asignadas por usuarios de Last.fm a la canción. Existen 19 541 etiquetas únicas. Reflejan percepciones y categorizaciones individuales de los oyentes.

2.2. Music4All - Onion [2]

Music4All-Onion organiza la información del contenido musical en capas (layers), que representan distintos tipos de características semánticas. La entidad de estudio principal continúa siendo la canción, pero se analiza desde múltiples perspectivas (audio, texto, imagen, interacción):

- **Capa de Audio (Audio Layer):** Contiene representaciones acústicas de las canciones extraídas mediante herramientas como *OpenSMILE*, *Essentia* y modelos de extracción de *i-vectors*. Incluye descriptores como patrones espectrales, MFCCs, jitter, shimmer, pitch, energía, entre otros. Ejemplos de archivos:
 - `id_mfcc_bow.tsv.bz2`: Bolsa de palabras (BoAW) construida a partir de coeficientes cepstrales en las frecuencias de Mel (MFCC).
 - `id_mfcc_stats.tsv.bz2`: Estadísticas agregadas de los MFCCs (media y matriz de covarianza aplanada).
 - `id_blf_spectral.tsv.bz2`: Patrones espectrales de baja fluctuación (BLF).
 - `id_blf_spectralcontrast.tsv.bz2`: Contraste espectral derivado del BLF.
 - `id_chroma_bow.tsv.bz2`: BoAW de características relacionadas al cromatismo tonal.
 - `id_ivec256.tsv.bz2`: Representación compacta basada en *i-vectors* calculados sobre MFCCs con 256 componentes de mezcla gaussiana.

Para mayor información revisar el **Cuadro: 2**

- **Capa Semántica de Letras (Extracted Metadata/Derived - EMD):** Incluye vectores semánticos derivados de las letras procesadas de las canciones. Se utilizan técnicas de procesamiento del lenguaje natural como *word2vec*, análisis de sentimientos y TF-IDF.

- `id_lyrics_word2vec.tsv.bz2`: Promedio de vectores *word2vec* preentrenados aplicados a las palabras de la letra.
- `id_vad_bow.tsv.bz2`: Representación BoW de puntuaciones de valencia, activación y dominancia extraídas con VADER.

Para mayor información revisar el **Cuadro: 3**

- **Capa de Contenido Generado por Usuarios (User Generated Content - UGC)**: Contiene información proveniente de las etiquetas (*tags*) asignadas por usuarios en plataformas como Last.fm, codificadas como vectores de TF-IDF.
 - `id_genres_tf-idf.tsv.bz2`: Representación TF-IDF de géneros asignados a cada canción.

Para mayor información revisar el **Cuadro: 4**

- **Capa Derivada de Contenido Visual (Derivative Content - DC)**: Basada en la extracción de características visuales de los videos musicales correspondientes. Estas se representan mediante embeddings generados por redes convolucionales preentrenadas como ResNet, VGG19 e Inception3.

Para mayor información revisar el **Cuadro: 5**

- **Eventos de Escucha (Interacciones Usuario–Canción)**: Conjunto de registros que reflejan las interacciones de los usuarios con las canciones. Contiene un total de **252 millones** de eventos de escucha, organizados en dos archivos principales:
 - `userid_trackid_timestamp`: Registros individuales de cada reproducción de una canción por parte de un usuario.
 - `user_id`: Identificador único asignado a cada usuario (provisto por Spotify).
 - `track_id`: Identificador único de la canción (coincide con el campo `id` en el resto de los datasets de Music4All).
 - `timestamp`: Marca temporal de la reproducción, en formato `yyyy-mm-dd hh:mm:ss`.
 - `userid_trackid_count`: Conteo agregado de cuántas veces un usuario ha reproducido cada canción.
 - `user_id`: Mismo identificador único del usuario.
 - `track_id`: Mismo identificador de canción.
 - `count`: Número total de reproducciones de la canción por ese usuario.

Para mayor información revisar el **Cuadro: 6**

2.3. Extrasensory [3]

La entidad de análisis es una **muestra de contexto**, dentro de un vector de características sensoriales y de estado asociado a un instante (1 min) y al usuario que lo generó. Cada muestra incluye: **[Revisar el Cuadro 7 para conocer la estructura del Dataset]**

- Metadatos de identificación
 - uuid – Identificador único de usuario (60 IDs)
 - timestamp – Marca temporal de la muestra
- Etiquetas de contexto
 - 50 columnas binarias, una por cada etiqueta (actividad o situación), con valor 1 si aplica, 0 si no.
 - Ejemplos: label_walking, label_sitting, label_meeting, label_home, etc
- Características sensoriales y de estado
 - Un total de 225 features agrupadas en seis grandes modalidades: acelerómetro, giroscopio, magnetómetro, audio, estado del teléfono y brújula del smartwatch.
 - Para cada sensor se extraen estadísticas temporales (media, mediana, desviación estándar, percentiles, ...)

Toda esta información se encuentra dividida de la siguiente forma en cada uno de los usuarios:

- **Formato de almacenamiento:**
 - Archivo principal: `extrasensory.csv`
 - Separador de columnas: `,` (CSV estándar)
 - Filas: 315 000 muestras (60 usuarios \times promedio de 5 000–6 000 muestras por usuario)
 - Columnas totales: 277
- **Metadatos de identificación:**
 - `user_id`
 - **Tipo:** text
 - **Descripción:** Identificador único asignado a cada usuario (60 IDs distintos).
 - `timestamp`
 - **Tipo:** datetime (formato `yyyy-mm-dd HH:MM:SS`)
 - **Descripción:** Fecha y hora exacta en que se tomó la muestra de sensores.
- **Etiquetas de contexto (ground truth):**
 - Cada etiqueta es una columna binaria (`boolean`) con valor ‘1’ si la condición aplica o ‘0’ en caso contrario.
 - Ejemplos de etiquetas (\approx 50 columnas):

- `label_walking` — Usuario caminando
 - `label_sitting` — Usuario sentado
 - `label_running` — Usuario corriendo
 - `label_lying_down` — Usuario acostado
 - `label_stairs_up` — Subiendo escaleras
 - `label_stairs_down` — Bajando escaleras
 - `label_stationary` — Usuario en reposo
 - `label_outside` — Usuario en espacio exterior
 - `label_house` — Usuario en casa
 - `label_library` — Usuario en biblioteca
 - ... (y más etiquetas relacionadas con actividad, ubicaciones y situaciones)
- **Características sensoriales y de estado del teléfono:** Los atributos se agrupan en seis modalidades principales. A continuación se describen, indicando el tipo de dato y la cantidad aproximada de columnas por grupo:

Acelerómetro (`accel_x`, `accel_y`, `accel_z`, etc.) — \approx **26 columnas**

- **Tipo de dato:** `double`
- **Descripción:** Estadísticos temporales (media, mediana, desviación estándar, percentiles) de las componentes X, Y y Z del acelerómetro, en unidades de m/s^2 .

Giroscopio (`gyro_x`, `gyro_y`, `gyro_z`, etc.) — \approx **26 columnas**

- **Tipo de dato:** `double`
- **Descripción:** Estadísticos temporales de las componentes X, Y y Z del giroscopio, en unidades de grados/s ($^\circ/\text{s}$).

Magnetómetro (`magnet_x`, `magnet_y`, `magnet_z`, etc.) — \approx **26 columnas**

- **Tipo de dato:** `double`
- **Descripción:** Estadísticos temporales de las componentes X, Y y Z del magnetómetro, en microteslas (μT).

Audio ambiente (`audio_rms`, `audio_peak`, `audio_centroid`, etc.) — \approx **13 columnas**

- **Tipo de dato:** `double`
- **Descripción:** Estadísticos del nivel sonoro del micrófono: RMS, pico, centroid, entre otros. Unidad aproximada: decibelios (dB).

Estado del teléfono (`battery_level`, `wifi_on`, `bluetooth_on`, `screen_on`, ...) — **7 columnas**

- `battery_level`: `double` (0.0–1.0), nivel de batería restante.
- `wifi_on`, `bluetooth_on`, `screen_on`: `boolean` (0 o 1) que indican si el Wi-Fi, Bluetooth o pantalla están activos.
- `airplane_mode`: `boolean`, indica si el dispositivo está en modo avión.

Brújula del smartwatch (`compass_mean`, `compass_std`, ...) — **6 columnas**

- **Tipo de dato:** double
- **Descripción:** Estadísticos de orientación (media y desviación estándar) en grados (0° – 360°).

2.4. Data Faltante

2.4.1. Ausencia de datos en Extrasensory

Es importante considerar que el dataset **Extrasensory** fue generado en condiciones reales y autónomas por los propios usuarios mediante la aplicación móvil *ExtraSensoryApp*. Esto introduce una serie de particularidades relacionadas con la calidad y completitud de los datos:

- **Etiquetas de contexto (label_...):** Las etiquetas de actividad, localización y situación fueron proporcionadas manualmente por los usuarios a lo largo del día. Debido a que estas anotaciones dependían de la iniciativa y disponibilidad del usuario, es común encontrar un alto porcentaje de valores faltantes (o ceros por omisión) en muchas de las columnas. Esta situación es esperable y no necesariamente indica mala calidad de datos, sino una dinámica natural en contextos no controlados.
- **Sensores del smartwatch:** Algunas características sensoriales (por ejemplo, las relacionadas con la brújula o sensores de muñeca) presentan una gran cantidad de datos ausentes. Esto se debe a que no todos los participantes contaban con un reloj inteligente emparejado con su teléfono durante la recolección de datos, por lo que estas variables no pudieron registrarse en todos los casos.
- **Variabilidad contextual entre usuarios:** Dado que los participantes provenían de distintos entornos y estilos de vida, no todas las clases o situaciones están igualmente representadas en la totalidad del dataset. Esta heterogeneidad puede conducir a un desbalance entre clases y variabilidad en la completitud por usuario.

En conjunto, estas características refuerzan la naturaleza “in-the-wild” del dataset, donde la falta de datos refleja condiciones reales de uso más que errores sistemáticos de adquisición, esto se contempla de mejor manera en la Figura 4, donde se observa el porcentaje de completitud promedio entre todas las columnas de todos los usuarios.

2.5. Granularidad Temporal

El dataset presenta una granularidad temporal de tipo **minuto a minuto**, ya que los registros de escucha de canciones se encuentran temporalmente codificados en formato `yyyy-mm-dd hh:mm:ss`. Esta precisión permite realizar análisis detallados sobre los hábitos de consumo musical de los usuarios a lo largo del tiempo.

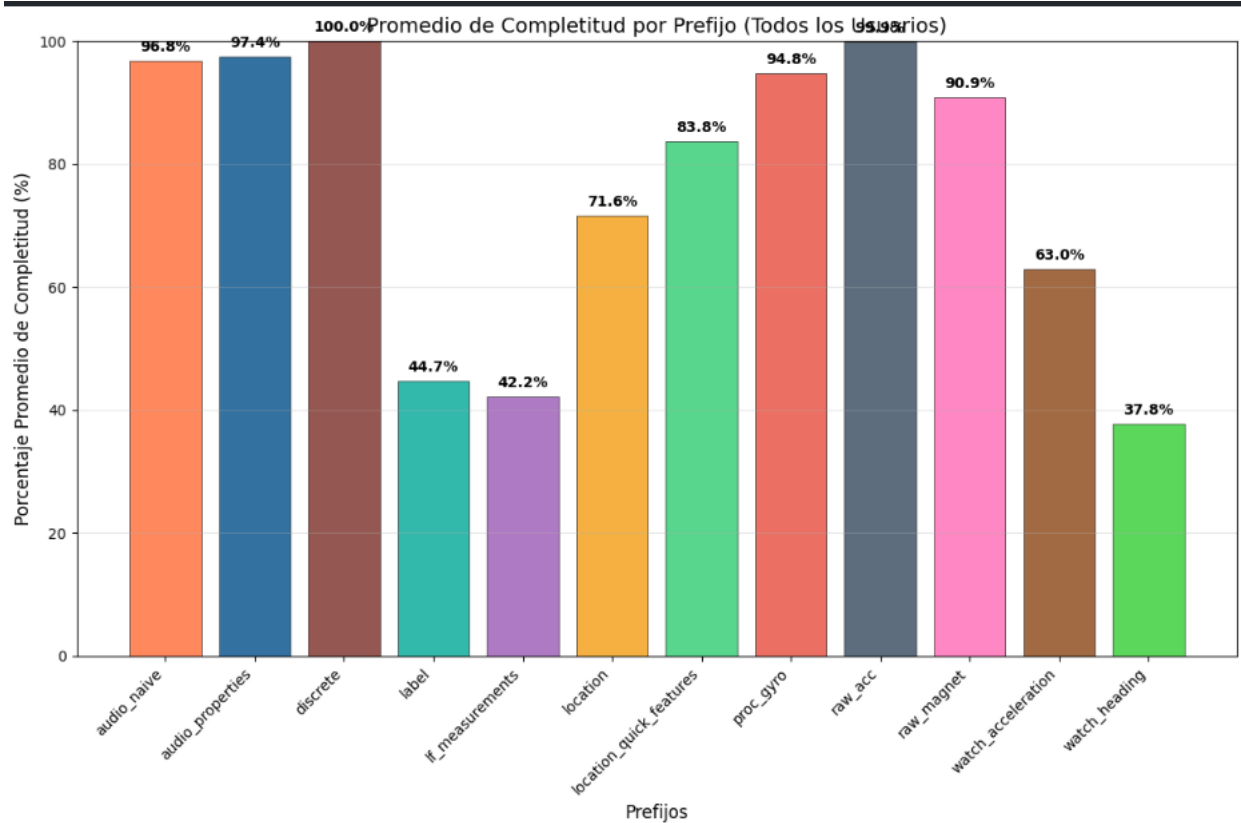


Figura 4: Porcentaje de datos completos entre todos los usuarios de extrasensory.

Se trata, por tanto, de un **problema dependiente del tiempo**, donde la dimensión temporal resulta crucial para comprender patrones de comportamiento. Aspectos como la variación de gustos musicales por estación del año, hora del día o nivel de actividad del usuario, pueden inferirse gracias a esta granularidad.

Además, la presencia de registros continuos a lo largo del año permite identificar **períodos de alta o baja actividad musical**, lo cual resulta relevante en contextos de recomendación o análisis de comportamiento. Por ejemplo, en la Figura 5 se observa un usuario con una distribución amplia de registros durante el año, así como una variabilidad notable en la **energía** de las canciones escuchadas en distintos momentos. Esto refuerza la importancia de considerar el eje temporal como una dimensión clave dentro del análisis.

En resumen, tanto la granularidad como la estructura temporal del dataset permiten modelar adecuadamente fenómenos secuenciales y contextuales relacionados con la escucha musical.

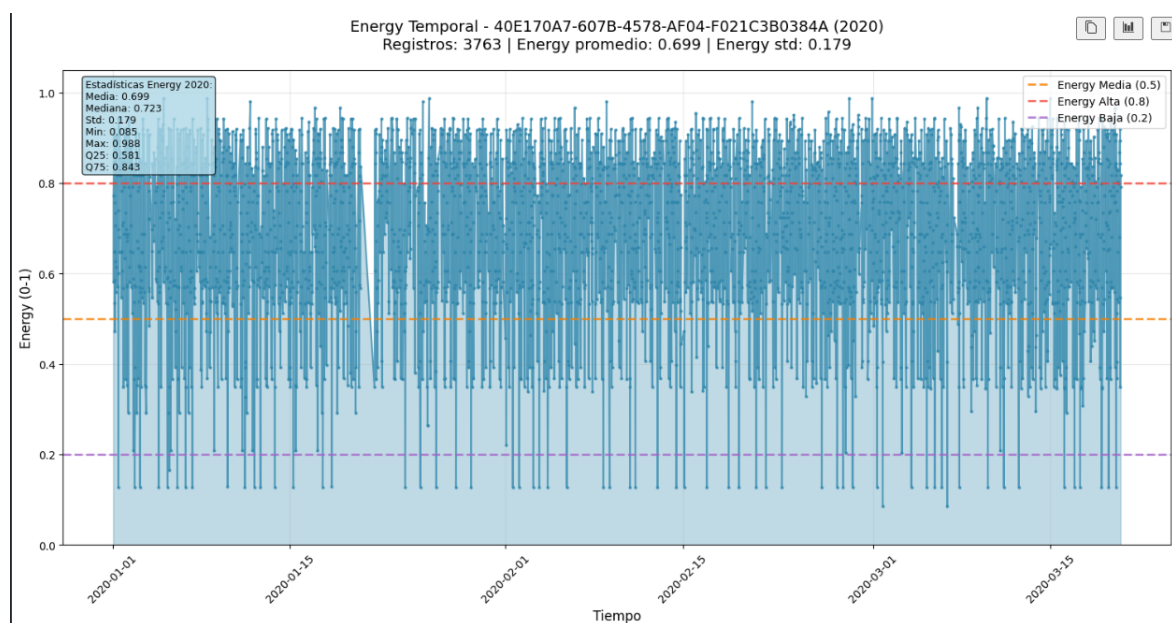


Figura 5: Comparativa temporal con la energía de las canciones escuchadas en 2020 por este usuario.

Cuadro 1: Descripción de atributos del conjunto de datos musicales

Atributo	Descripción	Tipo de dato	Rango/Valores posibles
id	Identificador único de la canción (16 caracteres)	Texto (string)	–
artist	Nombre del artista según Last.fm	Texto (string)	16,269 artistas únicos
song	Título de la canción	Texto (string)	–
lang	Idioma detectado de la letra	Texto (string)	46 idiomas únicos
spotifyid	Identificador de la canción en Spotify	Texto (string)	–
popularity	Nivel de popularidad en Spotify	Entero	0 a 100
album name	Nombre del álbum	Texto (string)	38,363 álbumes únicos
release	Año de publicación	Entero	Año (ej. 1960–2020)
danceability	Medida de qué tanailable es una canción	Decimal (float)	0.0 a 1.0
energy	Intensidad o actividad percibida en la canción	Decimal (float)	0.0 a 1.0
key	Tónica o clave principal (Pitch Class notation)	Entero	0 (C) a 11 (B)
mode	Modalidad (mayor o menor)	Binario (entero)	0 = menor, 1 = mayor
valence	Medida de positividad o felicidad	Decimal (float)	0.0 a 1.0
tempo	Tempo musical en beats por minuto	Decimal (float)	Depende de la canción (ej. 60–200 BPM)
genres	Lista de etiquetas de género asociadas a la canción	Lista de strings	853 géneros únicos
tags	Etiquetas generadas por usuarios en Last.fm	Lista de strings	19,541 etiquetas únicas

Cuadro 2: Descripción de archivos de características musicales

Archivo	Descripción técnica	Tipo
id_blf_*	Patrones espectrales (Band Energy, Spectral Flux) y patrones de fluctuación (BLF - Band-wise Local Feature) para análisis temporal-espectral	Audio BLF
id_chroma_bow.tsv.bz2	Representación Bag-of-Words (BoW) de características cromáticas (Chroma) que capturan la información tonal en 12 clases de altura	Audio OpenSMILE
id_compare_*.tsv.bz2	Estadísticas acústicas del conjunto ComPa-rE: MFCCs (Mel-Frequency Cepstral Coefficients), pitch fundamental, shimmer, jitter y parámetros espectrales	Audio OpenSMILE
id_emobase_*.tsv.bz2	Características acústicas relacionadas con emociones (EmoBase), incluyendo prosodia, calidad vocal y descriptores espectrales emocionales	Audio OpenSMILE
id_essentia.tsv.bz2	Descriptores musicales avanzados calculados con Essentia: características espectrales (centroide, flatness), rítmicas (BPM, onset rate) y tonales (key, scale)	Audio Essentia
id_mfcc_*.tsv.bz2	Coeficientes Cepstrales de Frecuencia Mel (MFCC) con sus estadísticas y representación Bag-of-Audio-Words (BoAW) para modelado tímbrico	Audio MFCC
id_ivec*.tsv.bz2	I-vectors (vectores de identidad) generados mediante Modelos de Mezcla Gaussiana (GMM) con diferentes configuraciones (256, 512 y 1024 componentes)	Audio MFCC

Atributo / Archivo	¿Qué representa?	Cómo se obtiene
processed_lyrics	Letras preprocesadas: normalizadas, lematizadas, sin anotaciones musicales	Procesamiento lingüístico y limpieza
id_lyrics_sentiment_functionals	Valores de valencia, activación y dominancia (emociones)	Léxico afectivo ANEW
id_lyrics_tf-idf	Representación TF-IDF (qué tan relevantes son las palabras en el corpus total)	Modelado de texto
id_lyrics_word2vec	Embeddings vectoriales de palabras, promediados por canción	Modelo Word2Vec preentrenado
id_vad_bow	Medidas emocionales según VADER, en formato "bag of words"	Léxico de sentimiento VADER

Cuadro 3: Descripción de los archivos y atributos utilizados en el análisis

Archivo	Descripción
id_genres_tf-idf.tsv	Representación TF-IDF de géneros musicales
id_tags_dict.tsv	Tags generados por usuarios de Last.fm con pesos asociados
id_tags_tf-idf.tsv2.bz2	Representación TF-IDF de tags

Cuadro 4: Archivos relacionados con metadatos generados por usuarios

Archivo	Descripción	Tipo (DC)
id_incp.tsv.bz2	Vectores de imágenes extraídas de videos usando Inception v3	Imagen
id_resnet.tsv.bz2	Vectores de imágenes extraídas con ResNet	Imagen
id_vgg19.tsv.bz2	Vectores de imágenes extraídas con VGG19	Imagen

Cuadro 5: Archivos con representaciones visuales extraídas de videos musicales

Archivo	userid_trackid_timestamp.tsv	userid_trackid_count.tsv
Descripción	Eventos individuales de escucha con timestamp	Conteo de reproducciones por usuario por canción
Nº Usuarios	119,140	119,140
Nº Canciones	56,512	56,512
Nº Registros	252,984,396	50,016,042

Cuadro 6: Resumen de archivos de historial de escucha

Modalidad	Cantidad	Descripción	Tipo de dato	Rango / Valores
Acelerómetro	26	Estadísticos (mean, std, median, ...) de aceleración en ejes X/Y/Z	Decimal (float)	Física (m/s^2)
Giroscopio	26	Estadísticos de velocidad angular en ejes X/Y/Z	Decimal (float)	Física ($^\circ/\text{s}$)
Magnetómetro	26	Estadísticos de campo magnético en ejes X/Y/Z	Decimal (float)	Física (μT)
Audio	13	Estadísticos de nivel sonoro (RMS, peak, centroid, ...)	Decimal (float)	Intensidad (varía según clip)
Estado teléfono	7	Indicadores de estado: nivel batería (0–1), Wi-Fi (on/off), Bluetooth (on/off), modo avión, etc.	Mix: float, binario	batería 0–1, flags 0/1
Brújula watch	6	Estadísticos de orientación (mean, std) extraídos del compass del smartwatch	Decimal (float)	Grados (0–360)
Etiquetas	~50	Self-reports binarios de actividades y situaciones	Binario (0/1)	0 = no aplica, 1 = aplica
Metadatos	2	uuid, timestamp	String, datetime	—

Cuadro 7: Descripción de modalidades, cantidad de variables y características del dataset

Referencias

- [1] I. A. Pegoraro Santana, F. Pinhelli, J. Donini, L. Catharin, R. B. Mangolin, Y. M. e. G. da Costa, V. Delisandra Feltrim, and M. A. Domingues, “Music4all: A new music database and its applications,” in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, Jul. 2020, p. 399–404. [Online]. Available: <http://dx.doi.org/10.1109/IWSSIP48289.2020.9145170>
- [2] M. Moscati, E. Parada-Cabaleiro, Y. Deldjoo, E. Zangerle, and M. Schedl, “Music4all-onion – a large-scale multi-faceted content-centric music recommendation dataset,” in *Proceedings of the 31st ACM International Conference on Information and Knowledge Management*, ser. CIKM ’22. ACM, Oct. 2022, p. 4339–4343. [Online]. Available: <http://dx.doi.org/10.1145/3511808.3557656>
- [3] Y. Vaizman, K. Ellis, G. Lanckriet, and N. Weibel, “ExtraSensory app,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, Apr. 2018, pp. 1–12.
- [4] —, “Extrasensory app: Data collection in-the-wild with rich user interface to self-report behavior,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’18. ACM, Apr. 2018, p. 1–12. [Online]. Available: <http://dx.doi.org/10.1145/3173574.3174128>