



INTELIGENCIA  
ARTIFICIAL

# SISTEMA DE RECOMENDACIÓN DE MÚSICA

JUAN ESTEBAN CAICEDO  
DANIELA OLARTE  
CARLOS PANTOJA  
CARLOS TAFURT

PRIMERA ENTREGA



**LOOKING FORWARD TO DEVELOPING A MULTIMEDIA PLATFORM THAT RECOMMENDS SONGS BASED ON THE TASTES OF THE USERS THEMSELVES, WE SEEK TO:**

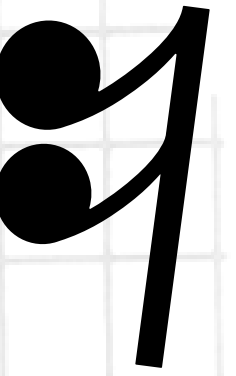
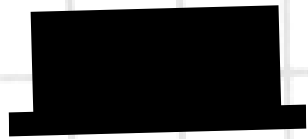
- DETERMINE WHAT TYPE OF ALGORITHM WOULD BE IDEAL FOR THIS TASK.
- USE DIFFERENT TECHNIQUES AND METHODS (K-MEANS CLUSTERING, THE CRISP-DM METHOD, ETC).

**THE MOST RELEVANT THING WAS:**

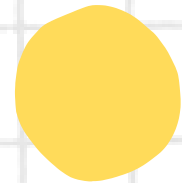
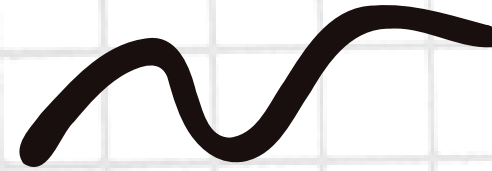
- FINDING THE CORRELATIONS OF THE DIFFERENT DATASETS SUCH AS THE HIGH CORRELATION BETWEEN THE POPULARITY AND THE YEAR OF LAUNCH OF THE SONG AS WELL AS THE POPULARITY AND THE ACOUSTICS OF THE SONGS.

**WE CAN ENSURE THAT THE PLATFORM WILL:**

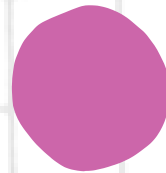
- IMPROVE THE USER EXPERIENCE
- IMPROVE THE RECOMMENDATION PROCESS
- REDUCE THE RATE OF ABANDONMENT OF PLATFORM USERS.



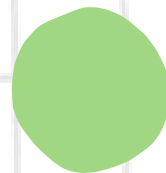
## PALABRAS CLAVES



Análisis exploratorio



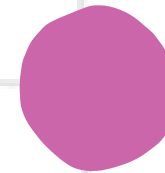
Segmentación



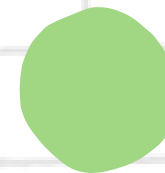
Recolección de datos



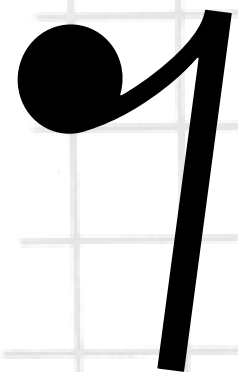
Procesamiento de datos



Limpieza de datos



Relevancia de la información





## INTRODUCCIÓN



Con el objetivo de recomendar canciones a los usuarios de una plataforma multimedia, tomando en cuenta **sus gustos** y los de **otros usuarios**, se busca implementar una aplicación que pueda cumplir con los objetivos anteriormente mencionados mediante el uso de **técnicas de inteligencia artificial**. Lo interesante de este proyecto se debe principalmente al hecho de poder **satisfacer los gustos** musicales de los usuarios, recomendando música que posiblemente les guste teniendo en consideración factores como lo son sus **artistas más escuchados**, su **género de mayor interés**, sus **canciones más oídas**, la **popularidad de la canción**, del **artista** y del **género**. Al poder acercar al usuario con nuevas canciones de su interés se logrará una mayor probabilidad que el usuario **continúe utilizando esta plataforma multimedia** y **la prefiera antes que a otras**.



## PREGUNTAS DE INTERÉS



1. ¿DE QUÉ MANERA SE PUEDE RECOMENDAR UNA CANCIÓN A UN USUARIO, CON BASE AL GÉNERO DE LA CANCIÓN?
2. ¿QUÉ ALGORITMO DE SEGMENTACIÓN ES EL MÁS ADECUADO PARA UN SISTEMA DE RECOMENDACIÓN DE MÚSICA?
3. ¿SE PUEDE RECOMENDAR UNA CANCIÓN A UN USUARIO, CON BASE A LA POPULARIDAD DE LA CANCIÓN, DEL GÉNERO Y DEL ARTISTA?





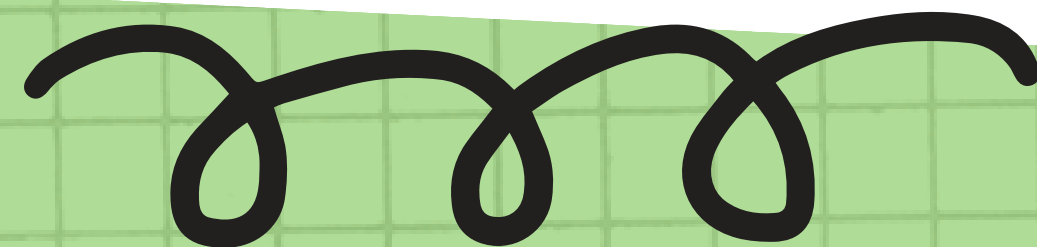
# MARCO TEÓRICO



1. **K-MEANS CLUSTERING ALGORITHM:** K-MEANS CLUSTERING ES UN ALGORITMO DE APRENDIZAJE NO SUPERVISADO QUE SE UTILIZA PARA RESOLVER LOS PROBLEMAS DE AGRUPACIÓN EN EL APRENDIZAJE AUTOMÁTICO O LA CIENCIA DE DATOS.
2. **CRISP-DM:** SE TRATA DE UN MODELO ESTÁNDAR ABIERTO DEL PROCESO QUE DESCRIBE LOS ENFOQUES COMUNES QUE UTILIZAN LOS EXPERTOS EN MINERÍA DE DATOS.
3. **SEGMENTACIÓN:** REFIERE AL ACTO DE SEGMENTAR DATOS DE ACUERDO A LAS NECESIDADES DE SU COMPAÑÍA, PARA ASÍ REFINAR SUS ANÁLISIS BASADOS EN UN CONTEXTO DEFINIDO, UTILIZANDO UNA HERRAMIENTA PARA ANÁLISIS DE CÁLCULOS VARIOS.



## MÉTRICAS DEL PROGRESO



- MÉTRICA 1. % DE DOCUMENTACIÓN HECHA
- MÉTRICA 2. % DE BACKEND REALIZADO
- MÉTRICA 3. % DE FRONTEND REALIZADO
- MÉTRICA 4. % DE PRUEBAS IMPLEMENTADAS

# Antecedentes

Los datos de referencia para el proyecto se tomarán de un dataset de Kaggle. Se escogió este dataset debido a que cuenta con la información adecuada para recomendar canciones a los usuarios, es decir, datos como el género de la canción, artista más escuchado, canción más escuchada, entre otras variables que pueden ser útiles para un sistema de recomendación de música.





## METODOLOGÍA

### CRISP-DM

Para el planteamiento del plan de este proyecto, se utilizó la metodología de CRISP-DM. A continuación se podrá observar los objetivos de la misma:





# DETERMINE BUSINESS OBJECTIVES



Se desea desarrollar una aplicación que permita hacer recomendaciones de canciones a los usuarios de una plataforma multimedia, tomando en cuenta sus gustos de este usuario y de otros usuarios.

# Ventajas y desventajas



## Ventajas:

- Menor gasto (contabilidad)

## Desventajas:

- No va acorde a los gustos de los usuarios.
- Cuando se lanza una nueva canción se debe ligar a otras de manera manual.
- No se mantiene un registro de la música favorita de cada usuario.

# BUSINESS OBJECTIVES

## Cuestiones comerciales:

- Mejorar la experiencia de usuario.
- Mejorar el proceso de recomendación.
- Reducir la tasa de abandono de los usuarios de la plataforma.

## Beneficios:

- Si se está conforme con el uso de la aplicación web debido a sus recomendaciones de música, los usuarios, gracias al voz a voz, provocarán un aumento sustancial de los mismos.
- Como consecuencia del anterior beneficio, se produciría una expansión del catálogo de música debido a que nuevos artistas quisieran agregar sus canciones a la aplicación web al ver un mercado potencial en esta.

## ASSESS SITUATION

### Inventory of resources

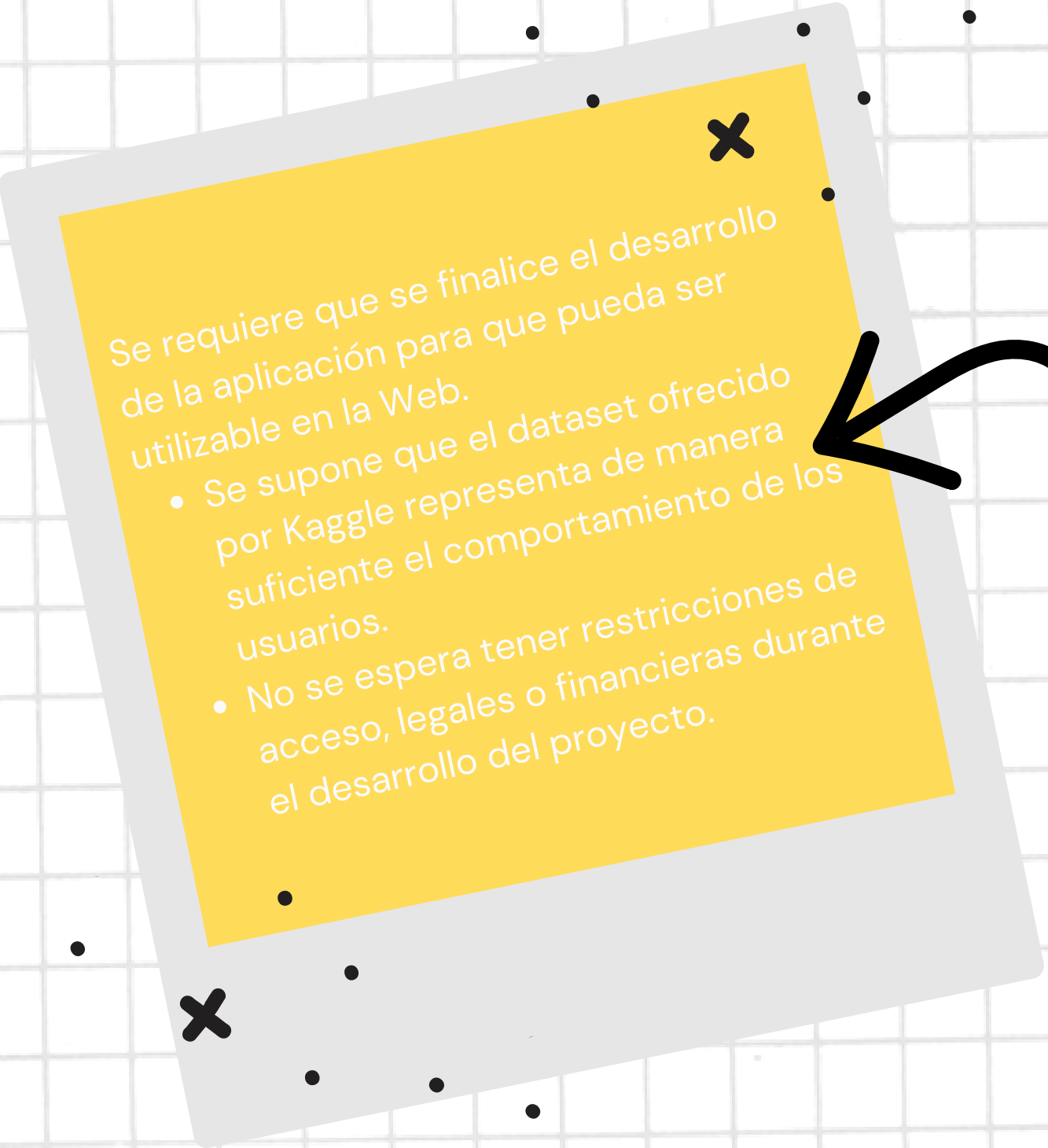
No se espera usar un hardware específico puesto que el despliegue será monolítico.

Los datos se almacenan en un dataset en forma de tabla. Se dispone del acceso directo a este dataset puesto que se pretende obtener de la plataforma Kaggle. Sin embargo, está pendiente por confirmarse si se podrá sacar el dataset de Kaggle.

## BUSINESS SUCCESS CRITERIA

- Recomendar música basada en el consumo musical individual.
- Recomendar música según el análisis de las calificaciones del usuario y de otros usuarios.





Se requiere que se finalice el desarrollo de la aplicación para que pueda ser utilizable en la Web.

- Se supone que el dataset ofrecido por Kaggle representa de manera suficiente el comportamiento de los usuarios.
- No se espera tener restricciones de acceso, legales o financieras durante el desarrollo del proyecto.

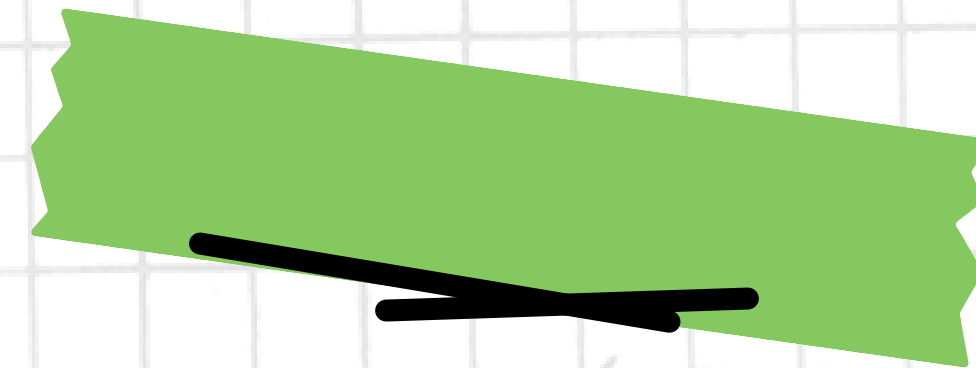
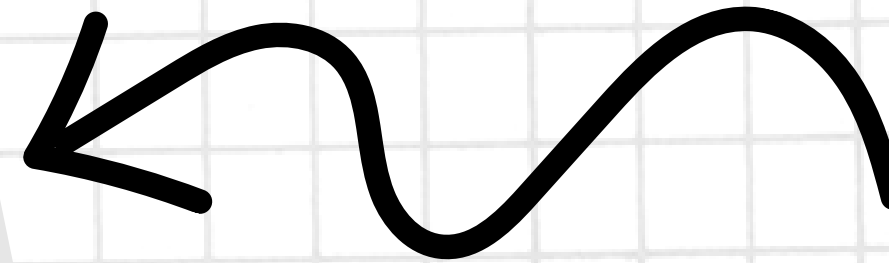
**Requirements,  
assumptions, and  
constraints**

## Costs and benefits

- No se esperan costos en la recolección de datos dado que, los datos necesarios para este proyecto se pueden recopilar en la propia plataforma de Kaggle.
- Se espera que los costos operativos sean bajos solo incluyendo el mantenimiento del software.
- Este proyecto puede ser beneficioso ya que a lo largo de su desarrollo se adquiere convicción en el manejo de datos y se da una mejor comprensión de ellos, lo que puede producir nuevas ideas donde se use la analítica.

# Determine Data Mining Goals

- Data mining goals
- El tipo de minería de datos que se va a utilizar es Segmentación.
- 3.2. Data mining success criteria
- La precisión como métrica de evaluación común para los problemas de clasificación. ✕
- Cómo se requiere que el proyecto sea utilizable, el despliegue satisfactorio de los resultados del modelo forma parte de los criterios de rendimiento.
- La matriz de confusión para saber cuántos valores clasificados fueron reales y no reales.



# Project plan

- Obtención y procesamiento de los datos
- construcción del modelo
- diseño del test
- evaluación del modelo
- despliegue

CONCLUSIONE

# CONCLUSIONES Y TRABAJO FUTURO



Se realizó un análisis visual para comprender mejor los datos de un dataset enfocado a la recomendación de música. Se aprendió principalmente como hacer un análisis exploratorio y de qué se que se compone el mismo. Así como también a identificar la relevancia de los datos para hacer uso de los mismos u optar por remover ciertos datos





ESO ES TODO

¡GRACIAS!

