



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



Minería de Datos

Evidencia de Aprendizaje

Ejercicio bases de datos

Docente: Mayra Cristina Berrones Reyes

Nombre: Gloria Nohemí Martínez Jiménez

Grupo:002

Matrícula: 1805800

12 de octubre de 2020

Google Play Store Apps

Esta base de datos muestra detalles de las aplicaciones en Google Play.

- *Objetivo:* Optimizar las recomendaciones que se hacen en Play Store de Google, se busca recomendar aplicaciones basadas en las categorías y géneros que sean más afines al usuario según su edad, tomando en cuenta sus aplicaciones favoritas.
- *Problema planteado:* Se suelen hacer recomendaciones de muchas aplicaciones con referencia a aplicaciones antes usadas, pero no se verifica que sean buenas aplicaciones y que a ese usuario le haya gustado la aplicación en la que se basó la recomendación, por lo que le guste o no esa aplicación le seguirán apareciendo recomendaciones de aplicaciones similares.
- *Solución:* Encontrar reglas de asociación basadas en las categorías y géneros de las aplicaciones, buscando que aplicaciones con buenas calificaciones se relacionan con otras aplicaciones de igual o mejor calificación, tomando en cuenta la edad del usuario y sus calificaciones en las aplicaciones.

Novel Corona Virus 2019 Dataset

Esta base de datos muestra información a nivel diario sobre el número de casos afectados, muertes y recuperación del nuevo coronavirus de 2019.

- *Objetivo:* Encontrar factores de riesgo de la enfermedad Coronavirus 2019 y así saber que medidas se deben de reforzar y cuales grupos de la población se verán mayormente afectados por la enfermedad, además de saber que grupos son más propensos a infección y distribución del virus.
- *Problema planteado:* La afectación del nuevo coronavirus 2019 empeora con ciertos factores, provocando la muerte del paciente y el contagio de personas a su alrededor.
- *Solución:* Realizar reglas de asociación entre las características de los pacientes, como su edad, sexo, lugar de residencia, y su progreso en la enfermedad, podemos ver si estos factores influyen en los días que estuvieron internados, si fallecieron, y conocer que grupos de la población tienen más riesgo de ser contagiados.

Wine Reviews

Esta base de datos muestra detalles de vinos como ubicación, bodega, precio y descripción.

- *Objetivo:* Crear un catálogo virtual con recomendaciones de diferentes tipos y marcas de vino para cadenas de restaurantes tomando en cuenta la clasificación y lugar de establecimiento del restaurante, además del nivel socio económico de los consumidores, para con ello saber con cuales vinos deben de surtir su inventario.
- *Problema planteado:* Los restaurantes buscan tener las mejores opciones para sus clientes, además de tener más seguridad de que su inventario se venderá, por lo que es necesario saber que productos serán mejor aceptados por sus consumidores.
- *Solución:* Encontrar reglas de asociación entre la ubicación del restaurante, el nivel socioeconómico de los clientes que asisten a ese restaurante con el precio y calidad del vino, además de buscar relaciones entre los demás alimentos que se tomaron en la orden para saber que vino recomendar.

Iris Species

Esta base de datos muestra tres especies de iris, así como algunas propiedades sobre cada flor.

- *Objetivo:* Optimizar una página de ventas en línea de arreglos florales para que cumpla con las expectativas exactas del cliente, realizar recomendaciones de posibles combinaciones en base a las características buscadas por la cliente y en base a esas características elegidas asignar un costo.
- *Problema planteado:* Clientes buscan arreglos florales con detalles específicos, como el color, longitud y altura de las flores por lo que se necesita un análisis de diferentes combinaciones posibles que se ajusten a las características buscadas.
- *Solución:* Crear un algoritmo que realice la clasificación de especies en base a sus características; como color, altura, anchura, etcétera; con ellas hacer asociaciones con posibles combinaciones para así recomendar posibles arreglos florales que le puedan gustar al cliente en base a las características deseadas.

Netflix Movies and TV Shows

Esta base de datos muestra detalles sobre programas de televisión y películas disponibles en Netflix a partir de 2019.

- *Objetivo:* Encontrar en qué clase de películas y series se deben de invertir más y cuando debe de hacerse, dependiendo de cuales son mejor aceptadas por el público según su lugar de residencia, edad, temporada del año y duración de la serie o película.
- *Problema planteado:* Para los creadores de contenido en Netflix es importante conocer que series y películas tendrán éxito en determinado tiempo y lugar del mundo, ya que crear o comprar los derechos de una serie o película es una inversión importante, también deben de saber que series o películas deben de ser retiradas de la aplicación o cuales deben de ponerse debido a la temporada.
- *Solución:* Desarrollar un algoritmo con clasificación de contenidos mejores aceptados para cada tipo de usuario en base a residencia, edad, y demás factores, tomando en cuenta su historial de vistas y duración en la aplicación, también se deben de buscar patrones secuenciales con base a la temporada del año.