



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

EJERCICIO BASES DE DATOS

NOMBRE: FERNANDO ISAI GONZALEZ CASTILLO

GRUPO: 002

MATRICULA: 1819011

MAESTRA: MAYRA CRISTINA BERRONES REYES

MATERIA: MINERÍA DE DATOS

FECHA: 13 DE OCTUBRE DEL 2020

a) Nombre de la base de datos: Google Play Store Apps.

Objetivo: Implementar una función que ayude al usuario a descargar la aplicación que necesita.

Problema planteado: En ocasiones tenemos muchas aplicaciones que tienen la misma función y descargamos la que primero nos aparece, cuando en realidad puede haber otras que podrían funcionar de mejor manera y no nos enteramos de que existen o tenemos que descargar varias aplicaciones hasta encontrar la indicada.

Solución: Desarrollar una función en el buscador de Google Play Store mediante la técnica de predicción, donde el usuario ingrese la categoría de la aplicación que esta buscando, para que quiere esta app, si esta dispuesto a pagar por ella o solo le interesan las aplicaciones gratuitas, el género de la aplicación, quien va a usar la aplicación (menor o mayor de edad) y si conoce el dato que introduzca la versión de su Android y el tamaño disponible en su celular, al final se le darán las aplicaciones que cumpla con todas o la mayoría de sus peticiones ordenadas por el número de personas que han instalado la aplicación o su calificación.

b) Nombre de la base de datos: Coronavirus.

Objetivos: Determinar el número de personas infectadas o fallecidas en un futuro y hacer que las personas sean conscientes sobre esta enfermedad.

Problema planteado: La enfermedad del Coronavirus ha sido muy controversial en el transcurso de este año, al principio las personas se cuidaban mucho, tomaban su sana distancia y utilizaban su cubrebocas, pero con el paso del tiempo la gente ya no tomó estas precauciones y los casos empezaron a aumentar en todo el mundo, por lo que debemos de hacer algo para que la gente se dé cuenta de que la enfermedad aún existe.

Solución: Dado los casos reportados de Coronavirus y los casos de personas fallecidas por esta enfermedad podemos realizar una regresión que se ajuste al modelo y se le puede mostrar al público para que vean a través de distintos escenarios como el número de casos está aumentando y lo que pasaría en un futuro si no tomamos precauciones, además de que podrán ver hasta cuando volveríamos a la normalidad.

c) Nombre de la base de datos: Criticas de Vinos.

Objetivos: Determinar el vino o los vinos más baratos con respecto a una característica en particular.

Problema planteado: Si no contamos con la cantidad de dinero necesaria para comprar un buen vino de calidad, buscamos el vino más barato y lo compramos sin importar sus características y puede no llegar a gustarnos o no cumplir con una de las cualidades que queríamos.

Solución: A través de la técnica de visualización podemos graficar el precio de los vinos con respecto a una característica que el usuario considere como primordial para su compra como por ejemplo el lugar de origen del vino (país, estado o región), un rango de calificaciones o alguna característica de la uva, es decir dado nuestro elemento primordial se realizará un grafico de dispersión y se le mostrarán los 3 vinos mas baratos con dicha cualidad.

d) Nombre de la base de datos: Clasificación de plantas.

Objetivos: Determinar a qué tipo de Iris (Setosa, Versicolor o Virginica) pertenece mi planta.

Problema planteado: En ocasiones no conocemos que tipo de planta tenemos por lo que terminamos por darle un cuidado inadecuado y genera que nuestra planta se marchite o en ocasiones si llegamos a consumir la planta sin conocer que planta es, podría ser nociva para nuestra salud.

Solución: Una solución sería realizar un clustering de nuestra base de datos con la información que tenemos de todas las plantas que fueron registradas como el Ancho y Largo del pétalo y sépalo, las cuales en teoría Iris Setosa formaría un cluster, Iris Versicolor otro cluster y Iris Virginica el último cluster, en donde cada cluster tiene características similares con los otros puntos que conforman ese mismo cluster, por lo que si el usuario ingresa las características de su planta, esta podrá pertenecer a un cluster y se podría determinar el tipo de Iris.

e) Nombre de la base de datos: Shows de Netflix.

Objetivos: Reducir el tiempo de búsqueda de un programa o película en Netflix.

Problema planteado: En ocasiones tomamos demasiado tiempo en elegir una película o un programa de televisión que cumpla con nuestros requisitos y si se le da al usuario algo de su elección que cumpla con sus requisitos, la persona verá Netflix y no pensará en cancelar su suscripción al no encontrar algo bueno.

Solución: Desarrollar un árbol de clasificación en forma de cuestionario o algún programa donde se le pregunte al usuario si desea ver una serie o película, su categoría favorita, su clasificación (en caso de que un niño vaya a ver Netflix), si desea ver un programa reciente o si tiene algún actor/actriz o director favorito y al final se le mostrarán las películas o programas que cumplan con todos los requisitos.