



**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN**  
**FACULTAD DE CIENCIAS FÍSICO-MATÉMICAS**



**Unidad de Aprendizaje:**

Minería de Datos

**Ejercicio Práctico Bases de Datos**

**Semestre Agosto 2020 – Enero 2021**

**Maestro(a):**

Mayra Cristina Berrones Reyes

**Alumno(a):**

Andrea López Solís 1822031

**Semestre: 7to.**

**Grupo: 003**

**Licenciatura en Actuaría**

Monterrey, Nuevo León, 14 de octubre de 2020

**Base de Datos:** Aplicaciones de Google Play Store.

**Objetivo:** Determinar cuáles son las aplicaciones mejor valoradas de Google Play Store por categoría, para hacer más ágil la búsqueda de recomendaciones y que así mismo evitar que los usuarios pierdan tiempo con aplicaciones de mala valoración.



**Problema planteado:** Cada vez son más los usuarios que se quejan de las malas recomendaciones por parte de la plataforma de Google Play Store, y debido a esto aumentan cada vez más las críticas negativas a las aplicaciones de baja valoración. Podríamos hacer listas de recomendaciones de las mejores aplicaciones yéndonos a cada categoría que maneja la plataforma de Play Store y de ahí ver cuáles son las aplicaciones más descargadas o buscadas por los usuarios; en base a la información recabada realizamos las listas donde se le muestre a los usuarios solo las aplicaciones mejor valoradas por categoría.

**Solución:** Desarrollar un programa o código que nos permita identificar cuáles son las aplicaciones más descargadas o buscadas por los usuarios, tomando en cuenta la categoría a la que pertenecen para hacer las recomendaciones acordes a éstas.

Desde el punto de vista de la minería de datos podríamos solucionar este problema apoyándonos de la técnica de Visualización de Datos, dado que esta nos permite observar la información mediante los distintos gráficos y nos permite apreciar de una forma más fácil cuales fueron las aplicaciones mejor valoradas por categoría, edad, género y mayor número de descargas.

**Base de Datos:** Novel Coronavirus 2019 Data.

**Objetivo:** Evaluar el comportamiento y expansión que tuvo la epidemia del COVID-19 en los principales países donde surgió para usar esa información en futuras epidemias o enfermedades nuevas que lleguen a surgir.



**Problema planteado:** Con el paso del tiempo cada vez son más las nuevas enfermedades que aparecen y a las que estamos expuestos, por ello es interesante estudiar el comportamiento de una nueva epidemia, dado que la información que lleguemos a recabar

de este suceso nos puede ser de gran utilidad para poder predecir el comportamiento, propagación, evolución y expansión de nuevas enfermedades que lleguen a surgir con el paso del tiempo y así mismo prepararnos ante la evolución de estas enfermedades que están apareciendo.

**Solución:** Mediante la técnica de predicción podemos desarrollar una herramienta que nos ayude a predecir el comportamiento de una nueva enfermedad, para ver cuánto tiempo tardaría en expandirse y cuánto tiempo le tomaría la sociedad o gobiernos tomar sus medidas de precaución para erradicar dicha epidemia.

Otra técnica que podríamos usar es la de visualización, ya que esta nos permite observar el número de contagios que se presentan por categorías y mediante esta información darnos una idea de lo que podríamos esperar en caso de que reaparezca el virus o se presente una nueva epidemia.

**Base de Datos:** Wine Reviews.

**Objetivo:** Determinar el país de donde provienen los mejores vinos, ver si todos los vinos de calidad proceden del mismo lugar o de zonas cercanas.

**Problema planteado:** Suponiendo que un comerciante o un conocedor de vinos quiere comprar vino de calidad para llevarlo y venderlo en su país o para él mismo, le sería de gran utilidad saber el país donde se elaboran los mejores vinos. Para ello consideramos el número de puntos WineEnthusiast de calificación que obtuvo cada vino y el lugar de procedencia.



**Solución:** Desarrollar un programa o herramienta que nos permita ver el lugar donde se elaboran los vinos de mejor calidad o los que son más valorados por los clientes.

Gracias a la técnica de visualización de datos podemos apoyarnos para observar mediante gráficos cuales son los vinos de mayor calificación y el lugar de procedencia de cada uno. Con la información recabada podemos ver si todos los vinos que son de buena calidad proceden del mismo lugar o de diferentes, esto facilitaría más la compra de vinos por parte de un comerciante, ya que este se enfocaría en ir a la zona donde estén los mejores vinos.

**Base de Datos:** Iris Species.

**Objetivo:** Determinar el tipo de especie al que pertenece una planta considerando las características señaladas en la base de datos.

**Problema planteado:** En caso de que una persona quiera buscar una planta, pero no sabe al tipo de especie al que pertenece dicha planta o su nombre; con solo ingresar las propiedades de la planta como sus medidas de sépalo y pétalo podría saber a qué especie pertenece y de que planta se trata.



**Solución:** Apoyándonos en la técnica de regresión lineal podemos observar cómo se relacionan las características de cada planta y en base a esta información podemos pasar a clasificar a que especie pertenece cada planta de acuerdo con sus características previamente ingresadas.

**Base de Datos:** Netflix Movies and TV Shows

**Objetivo:** Determinar que es lo que más ven los usuarios, si películas o series, para que Netflix se centre en invertir en lo que es más rentable y crear contenido de calidad que haga que los usuarios se interesen en ver.

**Problema planteado:** En la plataforma de Netflix podemos encontrar demasiado contenido nuevo que han estado agregando, donde casi la mayoría son series con duración de 1 temporada y de las cuales los usuarios han reaccionado no de manera positiva. Por ello sería de gran utilidad saber que quiere el usuario, cuáles son sus preferencias o gustos, para que Netflix como empresa apueste más a crear contenido enfocado en lo que los usuarios piden, y no tanto en solo crear contenido nuevo, sino también de calidad.

**Solución:** Con la técnica de visualización de datos podemos observar que es lo que más atrae al público si películas o shows de TV, y en base a esta información la empresa Netflix puede seleccionar en que proyectos debería de invertir para así atraer a más masas o mantener a los usuarios con los que ya cuenta.

