



UNIVERSIDAD AUTÓNOMA DE NUEVO  
LEÓN

FACULTAD DE CIENCIAS FÍSICO  
MATEMÁTICAS



# MINERÍA DE DATOS

Evidencia de Aprendizaje:  
Ejercicio práctico bases de datos

**Profra.** Mayra Cristina Berrones

**Alumna:** SABINA ALEJANDRA CASTILLO TRUJILLO

**Matrícula:** 1804029

L.A. Licenciatura en Actuaría

**FECHA:** 13 de octubre del 2020

**Nombre de la base de datos:** Google Play Store

**Objetivo:**

Establecer un precio adecuado y que la gente esté dispuesta a pagar para una app de diseño gráfico que está en desarrollo

**Problema planteado:**

Existen muchas apps destinadas al arte y diseño, sin embargo, su precio puede llegar a ser muy alto o parece ser que la app es gratis pero dentro de ella se requieren hacer muchas compras para acceder a ciertas herramientas o características

**Solución:**

Establecer UN solo precio para nuestra aplicación basado en el precio de las apps de este tipo con mejores ratings y comentarios. Se pudiera utilizar la técnica de regresión donde las variables sean las características de dicha app y nos de como respuesta un precio aceptable. La base de datos de igual manera nos pudiera ayudar a predecir entre qué grupos de edades tendrá mayor aceptación esta nueva aplicación

**Nombre de la base de datos:** Novel Corona Virus 2019 Dataset

**Objetivo:**

Poder identificar desde casa la presencia de corona virus con base en los síntomas que una persona llegue a padecer, sin necesidad de tener que ir a un hospital (en caso de que no sea grave)

**Problema planteado:**

El coronavirus es una enfermedad altamente contagiosa, por lo que la gente, al querer ir a un hospital a hacerse la prueba de covid, puede contraer el virus ahí mismo al convivir con demás gente posiblemente contagiada

**Solución:**

Con el uso de la técnica de reglas de asociación podremos identificar los síntomas más comunes, pero a su vez lograr diferenciar aquellos síntomas que vienen acompañados, de esta forma evitar que el coronavirus sea confundido con alguna otra enfermedad respiratoria y aislarse en la casa lo más pronto posible antes de poder contagiar a alguien

más. O en su defecto, saber identificar cuándo realmente es alguna otra enfermedad y no corona virus

**Nombre de la base de datos:** Wine Reviews

**Objetivo:**

Como dueña de una cadena de restaurantes en todo Estados Unidos me gustaría saber qué vinos implementar en cada una de mis sucursales ubicadas a lo largo del país

**Problema planteado:**

A pesar de que estamos hablando de un mismo país (Estados Unidos) existen diferencias, tanto en cultura como en gustos culinarios de la gente dentro de cada estado. Dicho esto, no puedo poner a la venta los mismos vinos en todas mis sucursales ya que existe la posibilidad de que no se vendan unos o que no tenga suficiente abastecimiento de otros

**Solución:**

Usar la técnica clustering de minería de datos, ya que esta se encarga de agrupar elementos en grupos con características similares. Es decir, agrupar a mis clientes de cada provincia de acuerdo con las características que éstos buscan en un vino para darle una buena calificación, ya sea que se fijen en el precio, la región de donde proviene, el tipo de uvas utilizadas, etc. Estos datos también me ayudarán a decidir con qué proveedores asociarme

**Nombre de la base de datos:** Iris Species

**Objetivo:**

Organizar un invernadero de tal manera que cada sección esté destinada a un solo tipo de Iris

**Problema planteado:**

Me dejaron a cargo de un invernadero y me dieron la indicación separar los tipos de Iris en secciones, ya que cada una requiere de cuidados y atenciones diferentes, sin embargo, no sé mucho de flores y se me dificulta diferenciar unas de otras

**Solución:**

Utilizar un árbol de decisión (específicamente uno de clasificación) que me vaya diciendo si la flor que estoy tratando de identificar es una iris-setosa, una iris-versicolor o una iris-virginica, dependiendo del largo y ancho de sus pétalos y de su tallo. Una vez identificadas y separadas, los iris podrán florecer de una manera óptima

**Nombre de la base de datos:** Netflix Movies and TV Shows

**Objetivo:**

Encontrar buenos actores/actrices que estén teniendo éxito y le gusten a la gente actualmente para la nueva película que se va a filmar

**Problema planteado:**

Un director quiere empezar a filmar una película, pero al casting no han llegado actores/actrices que sean adecuados para el papel, ya que el director está buscando actores/actrices con una carrera ya formada y buena reputación para que así la película se haga más famosa

**Solución:**

Ir encontrando patrones entre las películas más gustadas por los espectadores y el cast que conforma dicha filmación para identificar aquellos títulos que sean gustados gracias a la buena actuación de los integrantes y no tanto por la trama o el género de la película, de tal forma que los actores se puedan adaptar a esta nueva filmación. La técnica de minería de datos de patrones secuenciales sería muy útil para llegar a dicha solución