



UNIVERSIDAD AUTÓNOMA DE
NUEVO LEÓN
FACULTAD DE CIENCIAS FÍSICO
MATEMÁTICAS

Minería de Datos
Ejercicio Técnicas Bases de Datos

Maestra: Mayra Cristina Berrones Reyes

Alumno: Juan Alfredo Cantú Zavala

Matrícula: 1810736

Nombre: Google Play Store Apps

Objetivo: imaginemos que nosotros somos una empresa que da a sus empleados cierta aplicación que les ayuda a realizar su trabajo, y nosotros queremos poder tomar una decisión entre si pagar o no por una aplicación basado en el ranking de las aplicaciones.

Problema planteado: en caso de necesitar una aplicación que nos facilite cierta tarea queremos ver si realmente vale la pena invertir en una aplicación para la realización de dicha tarea porque para nosotros como empresa es importante reducir costos lo más posible y con esto aumentar los mas que se pueda la utilidad que generamos, entonces cada costo que tiene la empresa nos importa o es suficiente con una aplicación gratuita o si se pueden llegar a los mismos resultados con una aplicación gratuita.

Solución: Desarrollar una herramienta en la que la máquina identifique que tipo de aplicación necesitamos basado en características como el género, el peso o los requisitos que busquemos para nuestra aplicación, y ya identificadas esas características revisar si es de paga o es gratuita y ver sus respectivos rankings para ver si es mucha la diferencia entre ellos o si no la es y en base en eso tomar la decisión.

Nombre: Novel Corona Virus 2019

Objetivo: somos una organización de beneficencia cuya misión es brindar insumos a la primera línea de defensa contra el Coronavirus que se encuentra principalmente en el continente asiático intentando frenar la propagación de este virus y así dar mas tiempo a otras naciones para hacer sus planes de contingencia.

Problema planteado: los recursos que vamos a enviar son limitados, por esta razón queremos enviar el apoyo a la localidad mas afectada, o bien, localidad con más contagios registrados, además tomaremos en cuenta que sabemos que la población mas vulnerable son las personas de 65 años en adelante al ser más vulnerables sabemos que estas personas requieren más atención medica y materiales médicos que los jóvenes en la mayoría de los casos es por eso que decidimos que este factor lo consideraremos importante para tomar la decisión de a donde enviar los insumos.

Solución: necesitaríamos desarrollar una herramienta que nos permita primero aislar los datos que usaremos para tomar la decisión que son la localidad en donde fue registrado el paciente contagiado y la edad del paciente, una vez que las tengamos aisladas lo siguiente que haremos es contabilizar el número de casos por zona y edad, ya con estos datos aislados los acomodaremos de manera que nos permita ver fácilmente en donde hay mas pacientes mayores de 65 años contagiados de Coronavirus y ahí enviar las provisiones médicas.

Nombre: Wine Reviews

Objetivo: una empresa vinícola busca entrar en el mercado con un nuevo vino, y la nueva marca busca ganar reconocimiento y poder aumentar el estatus del producto contratando a un catador a que pruebe su producto y le otorgue una buena calificación a este.

Problema planteado: cuando una empresa introduce un producto nuevo en el mercado, la empresa debe convencer al cliente de que tiene un producto de calidad, que vale el precio que fijaron y que es mejor que el de los competidores (en caso de haber), esto se hace con publicidad y otras cosas, en esta industria donde existen catadores que otorgan una calificación a producto esta calificación puede darle mas prestigio al producto, generarle status y permitirle al productor ofrecer el producto mas caro justificado en el hecho de que un experto que evaluó su producto le puso una buena calificación.

Solución: para solucionar esto hay que desarrollar una herramienta que nos permita reconocer al catador que tenga mejor promedio de calificaciones, es decir, el catador que otorgue las calificaciones mas altas a los vinos ya que esto puede ser un indicador de que este critico es de cierta manera “benevolente” y así poder obtener la calificación mas alta posible al pedirle que valúe nuestro vino, lo que nos permitirá ganar cierto estatus y prestigio a nuestra marca que en mercados tan competitivos pueden marcar la diferencia entre un producto con altas y bajas ventas.

Nombre: Iris Species

Objetivo: supongamos que somos una florería que formará parte de un festival de parte de un festival de flores que se llevara acabo fuera de la ciudad, entonces la florería quiere saber si los Iris con el sépalo más ancho tienen los pétalos más anchos que las flores con sépalo más delgado.

Problema planteado: la florería tiene tiempo para plantar la cantidad de flores necesarias para el festival y como este evento es fuera de la ciudad las flores con pétalos mas gruesos son menos propensos a romperse durante la transportación que las flores con pétalos delgados, en caso de que sí estén relacionados plantar los Iris con el mismo método o semilla con las que fueron plantados los Iris con pétalo grueso.

Solución: esta situación la podríamos resolver desarrollando una hermanita que nos ayude a ver si existe una relación entre la anchura del sépalo y la del pétalo separando estos datos del resto y contando los casos en que se cumple que tienen un pétalo y sépalo ancho, si el número lo

consideran relevante, en base a esto se puede tomar la decisión de reproducir la planta de sépalo ancho o buscar otra opción.

Nombre: Netflix Movies and TV Shows

Objetivo: la casa productora de Netflix desea realizar una serie que dure más de una temporada y quiere decidir si darla más importancia a la elección del elenco de actores, o bien enfocarse en contratar a un director que ya haya hecho series de más de una temporada.

Problema planteado: Netflix quiere invertir en la producción de una nueva serie, pero para hacer esta inversión quieren tener una posibilidad de seguir la serie más de una temporada y para lograr esto buscan tomar una decisión con fundamentos de que importa mas entre elegir un director que tenga experiencia haciendo series que duren más de una temporada o bien en contratar actores que hayan participado en series de mas de una temporada, considerando los datos que ya se tienen de las series que hechas con anterioridad.

Solución: ya que quieren poder tomar una decisión basada en las series que ya hicieron hay que desarrollar una herramienta que logre identificar el nombre del director y de los actores en su respectiva columna y ya identificados los nombres ver cuales de ellos participaron en series cuya duración fue mayor a una temporada, con esta información ver cuales se repiten mas si los directores o los actores, aunque es evidente que habrá mas actores en cantidad lo que buscamos es ver cuantas series de mas de una temporada ha hecho cada uno.