



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

“Yo, como integrante de la comunidad estudiantil del Tecnológico de Monterrey, soy consciente de que la trampa y el engaño afectan mi dignidad como persona, mi aprendizaje y mi formación, por ello me comprometo a actuar honestamente, respetar y dar crédito al valor y esfuerzo con el que se elaboran las ideas propias, las de los compañeros y de los autores, así como asumir mi responsabilidad en la construcción de un ambiente de aprendizaje justo y confiable”

“Inteligencia artificial avanzada para la ciencia de datos I”

Análisis del contexto y la normatividad

Alumno:

Guillermo Romeo Cepeda Medina A01284015

Profesores:

Ivan Mauricio Amaya Contreras

Antonio Carlos Bento

Frumencio Olivas Alvarez

Blanca Rosa Ruiz Hernandez

Hugo Terashima Marín

Fecha de entrega: 29 de agosto de 2023

Introducción:

En una de las prácticas del módulo de estadística, utilizamos un dataset del menú de mc-donalds, este contiene varias características:

```
Index(['Category', 'Item', 'Serving Size', 'Calories', 'Calories from Fat', 'Total Fat', 'Total Fat (% Daily Value)', 'Saturated Fat', 'Saturated Fat (% Daily Value)', 'Trans Fat', 'Cholesterol', 'Cholesterol (% Daily Value)', 'Sodium', 'Sodium (% Daily Value)', 'Carbohydrates', 'Carbohydrates (% Daily Value)', 'Dietary Fiber', 'Dietary Fiber (% Daily Value)', 'Sugars', 'Protein', 'Vitamin A (% Daily Value)', 'Vitamin C (% Daily Value)', 'Calcium (% Daily Value)', 'Iron (% Daily Value)'], dtype='object')
```

Este dataset fue tomado de la plataforma kaggle, ellos mencionan que los datos fueron tomados directamente de la página web de mc-donalds. Sin embargo en la plataforma no se da crédito a la fuente ni se encuentra una licencia para usar los datos. No obstante en la página oficial de Mcdonalds Estados Unidos sí se encuentran estos datos por tipo de comida en el menú.

Aquí un ejemplo: [1]

Los datos del menú de macdonalds son de dominio público, debido a la ley que estableció el gobierno de Estados Unidos en mayo 7 de 2018, dónde establecen que los menús de restaurantes tienen como requerimiento dar a conocer el valor de sus calorías y demás valores nutrimentales en el menú. Por lo que mcdonalds sólo simplifica su acceso al publicarlos en su página web. No obstante, estos datos sólo aparecen en la web de Estados Unidos, en la web de México sólo aparece el menú con imagen y descripción de la comida, pero nada de los detalles nutricionales.

Según la definición de la UNESCO, la información de dominio público es definida cómo:

En la Recomendación sobre la promoción y el uso del plurilingüismo y el acceso universal al ciberespacio de la UNESCO figura la definición siguiente: "la expresión información de dominio público designa la información a la que el público puede acceder sin infringir ninguna disposición jurídica ni obligación alguna de confidencialidad. Por consiguiente, se refiere, por un lado, al conjunto de obras u objetos de derechos conexos que toda persona puede explotar sin autorización, por ejemplo debido a que no están protegidos en virtud de la legislación nacional o el derecho internacional, o a que el plazo de la protección ha expirado. Por otro lado, se refiere a los datos de carácter público y la información oficial producidos y difundidos voluntariamente por los gobiernos o las organizaciones internacionales".

Extraído de: [2]

Por lo que sí el acceso a la información no infringe ninguna ley o tiene obligación de confidencialidad, se puede decir que la información es de dominio público.

No obstante, es importante mencionar que los datos obtenidos del dataset de kaggle, fueron publicados hace 6 años, por lo que están sujetos a ser desactualizados y en el peor de los casos tener valores erróneos que dirijan a una mala interpretación y difamación de los mismos. Esto es algo muy grave y tiende a terminar en una demanda por parte del difamado y en México, con la ley de calumnia: [3] Puede terminar en un proceso penal. Existen muchos casos de demandas por uso incorrecto de datos como se puede observar en este reporte de Natassha Selvaraj: [4]

El uso erróneo de los datos del dataset de kaggle se puede representar en diferentes situaciones, por ejemplo si utilizo estos datos como una fuente de información oficial y publico un análisis que lleve a la conclusión de que hoy en día la comida de mcdonalds es dañina para la salud en México, estaría dañando a la compañía, debido a que estos datos ni si quiera son públicos en México y están desactualizados, además de que este análisis debería ser respaldado por algún profesional en nutrición para hacer una conclusión tan severa. Además de que si yo hago público un análisis que indique una correlación errónea de los datos, o una predicción falsa, el público será mal influenciado por mi análisis. En este caso la negligencia sería utilizar los datos de kaggle como reales y no tomar en cuenta su fecha, ni dar datos de error dentro del mismo dataset (Es decir, tal vez en la página de mcdonalds dice que sus hamburguesas tienen 300 calorías, pero se ha encontrado un error de +- 100 calorías en sus hamburguesas). Los datos deben ser tratados con cuidado, así como su interpretación.

[1]

<https://www.mcdonalds.com/us/en-us/product/as-featured-in-10-piece-mcnuggets-meal.html#accordion-c921f9207b-item-283bee7dbd>

[2]

https://www.oas.org/es/sla/ddi/docs/acceso_informacion_desarrollos_otros_UNESCO_orientacion.pdf

[3] <https://www.diputados.gob.mx/sedia/sia/spi/SAPI-ISS-12-12.pdf>

[4] <https://www.diputados.gob.mx/sedia/sia/spi/SAPI-ISS-12-12.pdf>Serving Size C