



**Tecnológico  
de Monterrey**

**Instituto Tecnológico y de Estudios  
Superiores de Monterrey**

**Campus Puebla**

**TC3006C. Inteligencia artificial avanzada para la ciencia de datos  
(Grupo 103)**

**Análisis del contexto y la normatividad**

Fernando Jimenez Pereyra

A01734609

15 de Junio de 2022

## **Licencia de los datos**

Los datos utilizados se encuentran publicados en kaggle bajo la licencia GPL2, la cual cubre la copia, distribución, y modificaciones. Si bien la licencia originalmente está orientada al software, será hará una interpretación general para el contenido, refiriéndonos a este como datos. Para entender cómo afecta la licencia en las cuestiones mencionadas podemos destacar los siguientes términos y condiciones de la licencia:

# 1 Copiar y distribuir: La licencia permite el copiar y distribuir libremente los datos.

# 2 Modificar: La licencia permite el modificar libremente los datos en cualquier proporción y distribuir dichas modificaciones libremente.

#5 Aceptación de la licencia: No se está obligado a aceptar la licencia, sin embargo al modificar o distribuir los datos se está aceptando los términos y condiciones de la licencia.

# 6 Redistribución: Cualquier redistribución de los datos debe de proveer de la misma licencia, sin la posibilidad de realizar cambios en la licencia o poner condiciones adicionales.

#11 Garantía: Los datos no presentan ninguna garantía, al menos de que se indique lo contrario.

## **Uso legítimo de los datos**

Los datos al estar distribuidos bajo la licencia GPL2 pude obtener una copia legal siendo plenamente consciente de que no puedo esperar una garantía de calidad de los mismos.

Al modificar los datos al momento de limpiarlo hago uso libre de mi derecho de modificarlo, como a su vez aceptó de forma definitiva la licencia por haber llevado dicha acción.

## Escenario de negligencia

Los dos escenarios principales que una persona podría caer en el mal uso de estos datos sería:

- La distribución de los datos originales o modificados sin respetar la licencia. Por ejemplo distribuir la licencia original o poner condiciones adicionales como un cobro por la distribución, lo que implicaría el incumplimiento de la licencia por la cual obtuvimos, y por lo tanto caeríamos en un acto ilegal, algo que únicamente se podría empeorar al tratar de suplantar la autoría original de los datos originales en el proceso.
- Actuar sin pensar en la garantía de la calidad de los datos. Si bien los datos son útiles para realizar una primera aproximación del modelo, el confiar ciegamente en los datos con la posibilidad de que se encuentren errados y por ende nuestro modelo no funcione adecuadamente en la realidad sería una conducta negligente, por lo que únicamente deberíamos de liberar una solución finalizada si pudiéramos cerciorarnos de la calidad de extracción de los datos utilizados, o probar nuestro modelo con datos de los cuales estemos consciente de la calidad de los mismos.

## Referencias

Free Software Foundation. (1991, junio). *GNU General Public License v2.0 - GNU Project*  
- *Free Software Foundation*. GNU Operating System. Recuperado 9 de septiembre de 2022, de <https://www.gnu.org/licenses/old-licenses/gpl-2.0.html>

*Fish market*. (2019, 13 junio). Kaggle. Recuperado 11 de septiembre de 2022, de <https://www.kaggle.com/datasets/aungpyaeap/fish-market>