

Instituto Tecnológico y de Estudios Superiores de Monterrey

Escuela de Ingeniería y Ciencias

Campus Monterrey

TC3006C.102: Inteligencia artificial avanzada para la ciencia de datos I

**Reporte Dataset Normatividad**

Alumno:

Josué Emmanuel Flores Mendoza - A00833132

Profesores:

Jesús Adrián Rodríguez Rocha

Lunes 9 de septiembre de 2024

## **Normativa Asociada al Uso de Datos**

El dataset "Heart Disease Prediction Dataset" se encuentra bajo la licencia MIT License. Esta licencia permite el uso, copia, modificación, publicación, distribución, sublicencia y venta de los datos de manera gratuita, siempre y cuando se incluya una copia de la licencia original en todas las distribuciones del software o derivados del mismo. Esto implica que se tiene la libertad de utilizar los datos con fines académicos, comerciales o de investigación sin restricciones significativas.

La normativa principal asociada al uso de este tipo de datos incluye:

- **Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP)** en México.
- **General Data Protection Regulation (GDPR)** en la Unión Europea.
- **Health Insurance Portability and Accountability Act (HIPAA)** en los Estados Unidos.

Dado que el dataset no especifica el origen geográfico de los individuos, he incluido la normativa mexicana, ya que es aplicable en el contexto local donde estoy ubicado, y las normativas de la Unión Europea y Estados Unidos, que son más generales e internacionales.

## **Medio de Utilización y Cumplimiento de Normativa**

### **Medio de Utilización:**

Estoy utilizando los datos para un proyecto de análisis de datos y predicción de enfermedades cardíacas, cuyo propósito es meramente educativo y de investigación. Los datos fueron obtenidos de una fuente pública en línea, donde el autor, Krish Ujeniya, ha proporcionado el dataset bajo la licencia MIT.

## Cumplimiento de Normativa:

Para asegurarme de no violar las normativas establecidas, he tomado las siguientes precauciones:

1. **Anonimización de los Datos:** Me aseguro de que los datos utilizados no contienen información que permita la identificación de personas de manera directa o indirecta. Esto se alinea con el principio de minimización de datos establecido en el GDPR y en la LFPDPPP, donde solo deben procesarse los datos necesarios para el propósito del análisis.
2. **Revisión de la Licencia MIT:** La licencia MIT proporciona permiso explícito para usar los datos sin restricciones, bajo la condición de que se mantenga la atribución del autor original. Este reporte incluye la información de la licencia y el autor original para cumplir con esta condición.
3. **Uso de Datos Públicos:** El dataset utilizado ha sido compartido en la plataforma Kaggle con la intención explícita de ser utilizado por investigadores, desarrolladores y educadores, lo que refuerza la idea de que su uso para análisis de datos no contraviene los términos de la licencia MIT.
4. **Ausencia de Datos Sensibles Individuales:** El dataset no contiene identificadores personales ni detalles que puedan usarse para identificar a individuos específicos, asegurando que no se está violando la privacidad ni las normativas de protección de datos personales.

## Documentos Oficiales de Referencia

- European Union. (2016). *General Data Protection Regulation (GDPR)*. Official Journal of the European Union.  
<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679>

- MIT License. (n.d.). *The MIT License (MIT)*. Open Source Initiative.  
<https://opensource.org/licenses/MIT>
- U.S. Department of Health and Human Services. (n.d.). *Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule*.  
<https://www.hhs.gov/hipaa/for-professionals/privacy/index.html>
- México. (2010). *Ley Federal de Protección de Datos Personales en Posesión de los Particulares*. Diario Oficial de la Federación.  
<https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>