



**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE  
MONTERREY**

*Inteligencia artificial avanzada para la ciencia de datos I (Gpo 102)*

**Momento de Retroalimentación: Reto  
Análisis del contexto y la normatividad.  
(Portafolio Análisis)**

Diego Loyo Villagrán A01571200

**9 de Septiembre de 2024**

## **1. Normativa asociada al tipo de datos utilizados (Iris dataset)**

El Iris dataset es un conjunto de datos de dominio público, ampliamente utilizado en estudios académicos, enseñanza y proyectos de machine learning. Este dataset no contiene información personal identificable (PII) ni datos sensibles. Por lo tanto, su uso no está regulado por normativas estrictas de protección de datos personales, como el Reglamento General de Protección de Datos (GDPR) en Europa o la Ley Federal de Protección de Datos Personales (LFPDPPP) en México.

### Documentación oficial:

El Iris dataset proviene del paquete de la biblioteca scikit-learn y está claramente identificado como datos públicos. La biblioteca menciona que estos datos son de libre uso, lo que significa que no están sujetos a restricciones legales como la recopilación de datos personales.

## **2. Medio de uso de los datos y medidas para asegurar cumplimiento normativo**

Se le dio uso a los datos a través de la biblioteca scikit-learn, una herramienta popular y reconocida en la comunidad de machine learning. La biblioteca scikit-learn garantiza que los datos incluidos son aptos para proyectos educativos y no contienen datos personales, por lo que para poder asegurarse de no violar ninguna normativa, se deben de checar las siguientes normativas:

GDPR (Artículo 2): Esta normativa no se aplica al conjunto de datos Iris, ya que no maneja datos personales.

Código ético en la ciencia de datos: Según las mejores prácticas en ciencia de datos, al analizar este dataset, se puede decir que se está utilizando un dataset ético y de dominio público, y no se recopila ni procesa información que pueda violar la privacidad de individuos.

### **3. Cumplimiento de la normativa en las herramientas utilizadas y prevención de sesgos éticos**

En este caso mi solución, al estar basada en regresión logística aplicada a datos públicos, no tiene violaciones éticas o legales relacionadas con la manipulación de información sensible. Sin embargo, para garantizar el cumplimiento normativo y ético en machine learning, es importante que:

Se evalúen posibles cambios, sesgos, o cosas extrañas en los algoritmos. Aunque el Iris dataset es sencillo y no representa a personas, en proyectos futuros que involucren datos más complejos será crucial realizar análisis de sesgo.

#### **Documentación oficial:**

El código ético de ACM (Association for Computing Machinery) para el desarrollo de soluciones tecnológicas establece que las herramientas deben ser transparentes, justas y no deben discriminar injustamente.

### **4. Escenarios de mal uso y posibles faltas éticas**

Si bien en este proyecto no hay implicaciones éticas graves, es importante considerar escenarios donde el mal uso de herramientas de machine learning pueda generar conflictos éticos. Estos podrían ser:

**Malicia:** Si se modifica la herramienta para clasificar información sensible de personas sin su consentimiento o si se emplea para propósitos malintencionados, como discriminación en áreas como empleo o finanzas.

**Negligencia:** La falta de monitoreo adecuado del desempeño del modelo podría resultar en errores significativos que afecten negativamente a quienes dependen de las predicciones generadas por la herramienta.