



**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)

Los datos que utilicé en mi proyecto corresponden al Iris dataset, un conjunto de datos de dominio público ampliamente usado en estudios académicos y enseñanza. Este dataset no contiene información personal identificable ni datos sensibles, por lo que no está regulado por normativas como el Reglamento General de Protección de Datos (GDPR) o la Ley Federal de Protección de Datos Personales (LFPDPPP). Al ser un dataset público, su uso es legal y no requiere ningún tipo de autorización especial.

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

Utilicé los datos a través de la biblioteca scikit-learn, la cual es ampliamente reconocida y respetada en la comunidad de machine learning. Esta biblioteca ofrece datasets etiquetados como públicos y aptos para fines educativos, lo que asegura que su uso no implica la violación de normativas de protección de datos personales.

Para garantizar que no estoy violando ninguna normativa, me baso en los siguientes principios:

- El Reglamento General de Protección de Datos (GDPR), que establece que la recopilación y tratamiento de datos personales debe estar sujeta a la regulación, no se aplica en este caso, ya que no hay datos sensibles ni personales.
- Además, sigo las mejores prácticas en ciencia de datos, asegurándome de que los datos utilizados son éticos y no violan la privacidad de ningún individuo.

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

Mi solución, basada en la regresión logística y aplicada a datos de dominio público, cumple con las normativas éticas en el ámbito de machine learning. Aunque el dataset de Iris no representa a personas, es importante estar consciente de posibles sesgos algorítmicos. Para futuros proyectos que involucren datos personales o más complejos, será crucial implementar análisis de sesgo para asegurar que las predicciones no discriminan injustamente.

El código ético de ACM (Association for Computing Machinery) establece que las herramientas tecnológicas deben ser transparentes y justas, y no deben discriminar de manera injusta. En mi caso, mi solución cumple con estos principios ya que el conjunto de datos y la herramienta utilizada están alineados con estas directrices.

4. Escenarios de mal uso y posibles faltas éticas

Aunque mi herramienta está diseñada para ser precisa y transparente, podrían surgir escenarios de mal uso en los que se incurra en faltas éticas. Por ejemplo:

- Malicia: Si la herramienta se adapta para clasificar personas o se utiliza de manera intencionada para discriminar a ciertos grupos, se podría caer en un uso malintencionado.
- Negligencia: Si no se monitorea adecuadamente el rendimiento del modelo y este se aplica en contextos donde las decisiones afecten a individuos sin las debidas validaciones, podrían ocurrir errores graves. Esto es particularmente importante en entornos críticos, como la medicina o las finanzas, donde los errores pueden tener consecuencias significativas.