

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

## Dataset del titanic

### Introducción

En este informe, se llevará a cabo un análisis del dataset utilizado en el módulo del reto para la competición de kaggle, "Titanic: Machine learning from disaster". El conjunto de datos elegido corresponde a información sobre los pasajeros a bordo del Titanic, disponible en la plataforma Kaggle. A continuación, se presentarán los hallazgos y el cumplimiento normativo relacionado con este proyecto.

### Normativa Asociada al Tipo de Datos

Este conjunto de datos del Titanic incluye información personal identificable, como los nombres de los pasajeros. En México, se aplican regulaciones de privacidad de datos personales, como la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP) <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf> . Por lo tanto, se deben cumplir estas regulaciones para garantizar el cumplimiento normativo.

### Cumplimiento Normativo

El proyecto cumple con los términos de uso de Kaggle, que establecen que los datos pueden utilizarse con fines educativos y de análisis, siempre que se respeten las condiciones establecidas por la plataforma. Los datos se utilizaron exclusivamente con fines educativos y de análisis, y se han seguido las políticas de uso de Kaggle. Además, se garantizó que no se alterara ni modificara el conjunto de datos original

### Cumplimiento de Normativas de la Industria y Ética

La solución desarrollada cumple con las mejores prácticas de la industria de la ciencia de datos y el aprendizaje automático, incluyendo las regulaciones específicas de privacidad de datos en México. Se han aplicado técnicas éticas y justas en el procesamiento de datos y en la implementación del modelo predictivo.

### Escenarios de Falta Ética Potencia

A pesar de las precauciones tomadas, es importante reconocer que si esta herramienta se utilizara con fines maliciosos o negligentes, podrían surgir problemas éticos. Por ejemplo, si se utilizara para tomar decisiones discriminatorias o para inferir información incorrecta sobre un grupo de personas en función de sus características demográficas, se estaría incurriendo en una falta ética.

### Conclusiones

En resumen, este proyecto de aprendizaje de máquina ha cumplido con las normativas relacionadas con el uso de datos disponibles en Kaggle. Se han seguido las políticas y

condiciones establecidas por la plataforma, y se han tomado medidas para garantizar la ética en el análisis de datos. Sin embargo, es fundamental recordar que la ética en la ciencia de datos es un aspecto crítico, y el uso inapropiado de los resultados puede tener consecuencias éticas negativas.

## Dataset Seleccionado

### Introducción

En este informe, se abordará el análisis de un dataset utilizado en el módulo de aprendizaje de máquina "Momento de Retroalimentación: Módulo 2 Implementación de una técnica de aprendizaje máquina sin el uso de un framework" [semana 3]. El dataset seleccionado corresponde a los datos del PIB per cápita de Canadá desde 1970 hasta 2020, obtenidos del sitio web del Banco Mundial. A continuación, se presentarán los hallazgos y el cumplimiento normativo relacionado con este proyecto.

### Normativa Asociada al Tipo de Datos

Los datos del PIB per cápita de Canadá son considerados datos económicos y financieros. La normativa relevante incluye regulaciones de privacidad de datos y acceso a información financiera gubernamental. En general, se debe cumplir con las regulaciones de privacidad de datos personales de Canadá, como la Ley de Privacidad de la Información Personal y Documentos Electrónicos (PIPEDA), pero estos datos no son datos personales, hablamos de datos de todo el país .

Según los términos de uso establecidos por el Banco Mundial, los datos publicados en su sitio web están disponibles para su uso libre, siempre y cuando no sean modificados ni alterados de ninguna manera. Sin embargo, es necesario reconocer que estos datos son de origen del Banco Mundial al hacer uso de ellos, y se debe hacer referencia explícita a la fuente original y otorgar el crédito correspondiente al Banco Mundial al publicar cualquier contenido en el que se utilicen tanto sus API como sus conjuntos de datos.

### Cumplimiento Normativo

El proyecto cumple con las normativas relacionadas con el uso de datos económicos y financieros según los términos de uso: <https://data.worldbank.org/summary-terms-of-use>. Los datos se utilizaron exclusivamente con fines educativos y de análisis, y no se incluyeron datos personales o información confidencial en el dataset. El acceso a los datos del Banco Mundial se obtuvo de manera legal y se siguió su política de uso de datos. Así mismo fueron referenciados tanto en el notebook como en el pdf.

### Cumplimiento de Normativas de la Industria y Ética

La solución desarrollada cumple con las mejores prácticas de la industria de la ciencia de datos y el aprendizaje automático. Se han implementado medidas para detectar y

mitigar sesgos en el modelo predictivo utilizado, como la selección de características y la selección de modelos.

### **Escenarios de Falta Ética Potencial**

A pesar de las precauciones tomadas, es importante señalar que si esta herramienta se utilizara con fines maliciosos o negligentes, podrían surgir problemas éticos. Por ejemplo, si se utilizara para establecer políticas económicas sin considerar otros factores o más tiempo de los datos ya que el lapso de tiempo es muy corto y la técnica de predicción usada para el reporte es muy simple “Regresión Lineal”, y se deberían de usar técnicas más robustas para predicciones tan importantes.

### **Conclusiones**

En general, este proyecto de aprendizaje de máquina ha cumplido con las normativas relacionadas con el uso de datos económicos y financieros. Se han tomado medidas para garantizar la ética en el análisis de datos y se han identificado posibles escenarios de falta ética. La transparencia y el cumplimiento normativo son esenciales en cualquier proyecto de ciencia de datos.

### **Referencias**

World Bank. (2023, 20 de julio). Resumen de los términos de uso. Recuperado de <https://data.worldbank.org/summary-terms-of-use>

Kaggle.(2023,1 de agosto). Submissions. Recuperado de <https://www.kaggle.com/competitions/titanic/submissions>.

México. (2010). Ley Federal de Protección de Datos Personales en Posesión de los Particulares. Recuperado de <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>