



INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE  
MONTERREY

INTELIGENCIA ARTIFICIAL AVANZADA PARA LA CIENCIA DE DATOS I

GRUPO 101

5 de septiembre de 2024

---

## Portafolio Análisis: Reporte de ética y normatividad

---

*Autor:*

Catherine Johanna Rojas Mendoza

A01798149

*Profesor:*

Edgar González Fernández

## Ética y Normatividad en el Análisis de Supervivencia del Titanic

Kaggle es una plataforma dedicada a la ciencia de datos y aprendizaje automático que brinda múltiples facilidades tanto para estudiantes como para profesionales. Permite el acceso a diversas bases de datos y la participación en competencias donde se resuelven problemas reales mediante modelos predictivos y técnicas avanzadas de análisis de datos. Los usuarios pueden crear y comparar soluciones, colaborando con una comunidad global que comparte recursos y conocimientos.

Además, Kaggle ofrece herramientas como notebooks interactivos, discusiones en foros, y un entorno de desarrollo basado en la nube, facilitando el aprendizaje continuo y el desarrollo profesional. Esto hace que sea un recurso valioso para quienes buscan mejorar sus habilidades en ciencia de datos y aplicar sus conocimientos en contextos prácticos, mientras participan en retos que permiten medir su desempeño frente a una comunidad global de expertos.

### Normatividad del reto

#### Terminos de Kaggle

1. **Propiedad de los datos:** Los usuarios deben tener los derechos necesarios sobre los datos que comparten en Kaggle, ya sea obteniendo permiso del propietario original o usando datos de dominio público (Kaggle, 2024).
2. **Permiso de publicación:** Los datos compartidos en Kaggle deben poder ser distribuidos públicamente y no infringir derechos de propiedad intelectual ni acuerdos de confidencialidad. Los usuarios son responsables de asegurar el cumplimiento de estas normas (Kaggle, 2024).
3. **Cumplimiento con GDPR:** Si los datos contienen información personal de residentes de la Unión Europea, deben cumplir con el GDPR. Los datos personales deben estar anonimizados o contar con consentimiento explícito (Kaggle, 2024).
4. **Normas de uso aceptable:** No está permitido compartir datos que sean ofensivos, dañinos o ilegales, ni aquellos obtenidos de manera no ética (Kaggle, 2024).
5. **Derecho de eliminación:** Kaggle puede eliminar datos que infrinjan sus términos de uso, violen leyes o derechos de autor (Kaggle, 2024).

#### Normas éticas y de cumplimiento en la competencia del Titanic

1. **Uso responsable de los datos:** Los datos del concurso solo pueden usarse con fines no comerciales, como investigación académica o para participar en la competencia (Kaggle, n.d.).

2. **Restricción en el uso de cuentas:** Solo se permite usar una cuenta por participante. El uso de múltiples cuentas conlleva descalificación (Kaggle, n.d.).
3. **Prohibición de compartir código en privado:** El código solo puede compartirse públicamente en los foros, y no de manera privada entre equipos, salvo que haya una fusión de equipos (Kaggle, n.d.).
4. **Garantía de uso ético de herramientas automáticas:** Se permite el uso de herramientas automáticas de aprendizaje automático, siempre que los participantes cuenten con las licencias necesarias (Kaggle, n.d.).
5. **Cumplimiento con restricciones geopolíticas:** Los participantes de países sancionados, como Cuba, Irán o Corea del Norte, no pueden participar, y deben cumplir con las leyes locales (Kaggle, n.d.).

### **Cumplimiento de Leyes, Normas y Principios Éticos**

La primera entrega, titulada *El Propósito Oculto de la Propuesta de Valor y sus Implicaciones Éticas en el Análisis de Supervivencia del Titanic*, así como la exploración y desarrollo de diversos modelos predictivos basados en los datos históricos del Titanic, destaca la importancia de atender las consideraciones éticas y legales relevantes para garantizar un análisis adecuado en el contexto actual.

La solución propuesta en la primera entrega, es el desarrollo de modelos predictivos utilizando datos históricos del Titanic de manera ética y responsable, siguiendo principios clave como la transparencia, el respeto a la dignidad humana, la justicia y la no discriminación. Se subraya la necesidad de que los modelos estén alineados con fines educativos o de investigación, evitando su uso comercial o trivial. Además, se enfatiza la importancia de aplicar los principios de minimización de datos y transparencia, respetando la memoria de las víctimas y evitando deshumanizar la tragedia al convertirla en simples estadísticas.

A pesar de esto, hay otros elementos éticos y legales que también deben considerarse, como el cumplimiento de leyes de protección de datos y el uso justo de la información, para asegurar que el tratamiento de estos datos históricos sea riguroso y respetuoso.

### **Sesgo en los perfiles de riesgo y toma de decisiones**

Los modelos predictivos pueden verse afectados por factores como la clase social, el género o la edad, lo que podría perpetuar desigualdades. Es clave aplicar principios de justicia y no discriminación para evitar estos sesgos y asegurar un análisis ético (Barocas et al., 2019).

### **Tratamiento de datos confidenciales**

Aunque los datos del Titanic no están regulados por leyes actuales como el GDPR, es importante seguir principios de minimización de datos y transparencia (GDPR, 2016).

### **Transparencia en la propuesta de valor**

Los objetivos al usar datos históricos deben ser claros y alinearse con fines educativos o de investigación, evitando su uso comercial que deshumanice a los involucrados (Breidbach & Maglio, 2020).

### **Normatividad aplicable**

Aunque los datos del Titanic no están sujetos a normativas como el GDPR, es recomendable seguir principios de protección de datos para promover un uso ético y respetuoso (Future of Privacy Forum, 2018).

El desarrollo de modelos predictivos con datos del Titanic plantea desafíos éticos que requieren una gestión cuidadosa. Cumplir con normas como la minimización de datos, la transparencia y la mitigación de sesgos es esencial para asegurar que estos modelos aporten de manera positiva. Es importante evitar trivializar experiencias humanas complejas y respetar la memoria de las víctimas, promoviendo un uso responsable y respetuoso de la tecnología.

### **Conclusión**

La solución propuesta para el análisis de la supervivencia en el Titanic cumple con las leyes, normas y principios éticos establecidos tanto en la industria de ciencia de datos como en el contexto específico del reto. En primer lugar, el uso de datos históricos se realiza con fines educativos y de investigación, lo que garantiza que no se exploten de manera inapropiada los datos ni se utilicen con fines comerciales o triviales. Este enfoque respeta el principio de transparencia, asegurando que el propósito del modelo sea claro y alineado con objetivos legítimos.

Parte de la normatividad incluye la prohibición de compartir datos o código de manera privada, la restricción a una cuenta por participante, y la aplicación de normas internacionales de competencia justa. También se destaca el cumplimiento con las leyes aplicables en cuanto a participación, como las regulaciones geopolíticas que prohíben la participación de residentes de ciertos países.

## Referencias

- Kaggle. (2024). Terms of Use. Retrieved from <https://www.kaggle.com/terms>
- Kaggle. (n.d.). Titanic: Machine Learning from Disaster Rules. Retrieved from <https://www.kaggle.com/competitions/titanic/rules>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning. Available online at <http://fairmlbook.org>.
- Reglamento General de Protección de Datos (GDPR) (2016). This regulation outlines the legal framework for the processing of personal data within the European Union. Available at <https://gdpr-info.eu>.
- Breidbach, C. F., & Maglio, P. (2020). Accountable algorithms? The ethical implications of data-driven business models. *Journal of Service Management*, *31*(3), 581-597.  
<https://doi.org/10.1108/JOSM-05-2019-0140>
- Future of Privacy Forum (2018). Best Practices for Privacy and Data Protection in Machine Learning. Available at <https://fpf.org>.