



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

**Inteligencia artificial avanzada para la ciencia de datos I
TC3006C.101**

Momento de Retroalimentación: Análisis del contexto y la normatividad

Marcelo Márquez A01720588

28 ago 2023

Para este análisis contextual y normativo, se ha elegido el primer entregable del Módulo 1 de Estadística: **Técnicas de Procesamiento de Datos para el Análisis Estadístico y la Construcción de Modelos**. El objetivo de este proyecto consistió en examinar un conjunto de datos que contenía información detallada sobre vehículos competitivos, incluyendo sus precios, con la finalidad de identificar los principales factores que influyen en la determinación de su precio de mercado.

1. Investiga cual es la normativa asociada al tipo de datos que utilizaste.

Mientras que los datos utilizados conforma con las leyes y regulaciones como la **Ley General de Protección de Datos Personales**, es igualmente crucial tener en consideración los estándares industriales relacionados con el análisis de datos y usos de bases de datos ajenos incluyendo o no datos personales.

Es imperativo comprender y respetar las normativas de propiedad intelectual como **Ley Federal del Derecho de Autor** y la **Ley de la Propiedad Industrial en México** (DOF, 2017). Estas leyes protegen los derechos de los propietarios de contenidos creativos y de la propiedad intelectual relacionada, lo que significa que cualquier información sujeta a derechos de autor o propiedad industrial contenida en el dataset debe ser tratada con el debido respeto y cumplimiento de las disposiciones legales correspondientes.

En el contexto del procesamiento de datos y su análisis estadístico, resulta esencial seguir los procedimientos adecuados y cumplir con estándares reconocidos, como el ISO/IEC 20000 (ISO, 2018), al emplear datos provenientes de empresas en competencia. Esto cobra especial relevancia teniendo en cuenta que el objetivo último de dicho análisis es su aplicación en el sector del mercado automovilístico.

2. Explica por qué medio utilizaste los datos y como te estas asegurando de no violar la normativa establecida. Respalda esto con los documentos oficiales en los que te estás basando y cítalos.

Los datos fueron proveídos como recurso dentro de la materia. Debido al medio por el cual fueron adquiridos los datos, se toma como supuesto que estos tienen una licencia de “uso libre” o *Open Database license*.

Esta licencia permite que uno comparta, modifique y use la base de datos libremente de manera que, al no estar lucrando del uso de la base de datos o previniendo que otras personas tengan acceso a la misma, me estoy asegurando de no actuar en contra de los lineamientos establecidos por la licencia (Open Knowledge foundation, s.f.).

De acuerdo a la Ley General de Datos Personales, se define *Fuentes de Acceso Público* de la siguiente manera:

“Aquellas bases de datos, sistemas o archivos que por disposición de la ley puedan ser consultadas **públicamente** cuando no exista impedimento por una norma limitativa y sin más exigencia que, en su caso, el pago de una contraprestación, tarifa o contribución. No se considerará fuente de acceso público cuando la información contenida en la misma **sea obtenida o tenga una procedencia ilícita**, conforme a las disposiciones establecidas por la presente Ley y demás normativa aplicable”.

(DOF 26-01 2017 pg. 4 de 52)

De acuerdo al **Artículo 22** de la Ley citada anteriormente, “El **responsable no estará obligado a recabar el consentimiento del titular para el tratamiento de sus datos personales** en los siguientes casos, **VIII. Cuando los datos personales figuren en fuentes de acceso público**. A pesar del hecho que la base de datos utilizada no contiene **Datos personales** ni **Datos personales sensibles**, conforme a las definiciones establecidas en la Ley, se toman en cuenta durante el análisis apegar a las disposiciones que la misma prevé.

El ISO/IEC 20000, un estándar para la gestión de servicios de tecnología de la información (TI), puede ser aplicable en el contexto académico. Ayuda a garantizar que los datos se utilicen de manera correcta y que se mantenga la integridad de la información. Esto es esencial para prevenir el uso indebido de los datos y asegurarse de que no se realicen publicaciones que mal informen o violen los derechos de terceros en el entorno académico. En el contexto del reto, el estándar ISO/IEC 20000 establece lineamientos para el mejor uso de los datos en una empresa. De manera que, con base al reto, si la empresa China quiere entrar a un

nuevo mercado se debe apegar a lo establecido en el estándar para cumplir con la normatividad de la base de datos.

Sin embargo, tomando en cuenta que la empresa China quiere entrar al mercado estadounidense, se toma en cuenta que no hay leyes de propiedad intelectual que protejan las bases de datos (WIPO, s.f.).

3. Una vez analizado el dataset ahora hay que analizar la herramienta, explica cómo tu solución cumple con la normativa establecida en la misma industria y no incurre en ningún sesgo ético.

Hay varios factores que verifican el cumplimiento de de las normativas del uso, modificación y análisis de los datos utilizados. La solución proporciona una descripción detallada de la metodología utilizada para analizar los datos de la industria automovilística. Esto incluye la selección de variables, técnicas de análisis estadístico y modelos utilizados.

La transparencia garantiza que los procesos sean claros y comprensibles, lo que contribuye a evitar sesgos ocultos. Otro factor es el uso de las variables seleccionadas y su significancia a la industria automovilística en cuanto a sus precios, el seleccionar variables que tengan un menor impacto en cuanto a sus precios puede causar una representación defectuosa que pueda afectar no solo al cliente sino a la misma empresa.

4. Analiza bajo qué escenarios se podría incurrir en una falta ética si tu herramienta es mal usada, contempla casos como malicia o negligencia.

En el caso de la base de datos con información de automóviles, se pudiera hacer mal uso de la misma para **difamar una compañía automovilística**, manipular a consumidores hacia realizar una compra con base en conclusiones obtenidas del análisis y se pudieran publicar los datos sin darle crédito al autor.

Desde el punto de vista de la falta ética, en el contexto del objetivo del análisis (identificar las características principales para predecir el precio de un automóvil), se podría manipular intencionadamente los datos con el propósito de generar un análisis defectuoso que justifique un aumento injustificado en el precio de los automóviles para los clientes. Esto perjudica a los consumidores al inflar los costos sin una base legítima.

Referencias

Diario Oficial de la Federación. (2017). Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados. Cámara de Diputados Del H. Congreso de la Unión.

Recuperado de: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGPDPPSO.pdf>

ISO. (2018). ISO/IEC 20000. Recuperado de <https://www.iso.org/standard/70636.html>

Open Knowledge Foundation. (s.f.) Open Data Commons Open Database Licence (ODbl)v1.0. Recuperado de: <https://opendatacommons.org/licenses/odbl/1-0/>

WIPO. (s.f.). Protection of Non-Original Databases. Recuperado de: <https://www.wipo.int/copyright/en/activities/databases.html>