



Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey

Momento de Retroalimentación: Análisis del contexto y la normatividad.

TC3004B.104 Inteligencia Artificial Avanzada para la Ciencia de Datos I

Profesores:

Ivan Mauricio Amaya Contreras

Blanca Rosa Ruiz Hernandez

Antonio Carlos Bento

Frumencio Olivas Alvarez

Hugo Terashima Marín

Julian Lawrence Gil Soares - A00832272

28 de Agosto de 2023

Dataset: <https://archive.ics.uci.edu/dataset/186/wine+quality>

Introducción

Este dataset es extraído de la base de datos de University of California Irvine (UCI) de parte de su base de datos de datasets para problemas de regresión y clasificación. El dataset consiste en 12 variables que se pueden utilizar en un modelo de clasificación o de regresión para determinar la calidad del vino que se produce entre vino tinto y vino blanco. Para nuestra entrega anterior implementamos un modelo de regresión, en mi caso de regresión lineal, para buscar qué clases de predicciones podríamos realizar acerca de la calidad de cada uno de los vinos.

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

Yo utilice los datos de la base de datos de uci, estos datasets son regulados por Creative Commons Attribution 4.0 International (CC BY 4.0) que son derechos de autor que regulan las maneras en las que los usuarios podemos modificar y distribuir los datos dentro del dataset. Esta normativa funciona como protección a la propiedad intelectual de quienes realizan las bases de datos y garantiza que el usuario tenga acceso a una libre navegación sin barreras que llegaran a incurrir en un abuso de confianza sobre la plataforma.

2. Explica porque 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.

Para el manejo de los datos utilice python y su librería pandas estas son herramientas gratuitas y de open source y los proyectos creados en ellos no son considerados como propietarios del autor si no como herramientas funcionales. Para cumplir con las normas establecidas por la licencia nos aseguramos de atribuir correctamente a los autores tal como se establece en la sección 3 del CC BY 4.0.

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.

Como se mencionó anteriormente, la normativa utilizada fue Creative Commons Attribution 4.0 International en la que se desarrollan mecanismos por los que se otorgan licencias de atribución para uso de prácticas aceptables.

Particularmente, la solución al problema planteado no incurre con sesgos étnicos a causa de la naturaleza de la misma normativa utilizada para el uso de la base de datos en la que se trabajó, puesto que no se incurrió en medios con fines lucrativos sino de calidad de diferentes tipos de vino usando una predicción basada en datos particulares.

El posible sesgo que podría llegar a ser aplicable dependerá de un mal uso de la información proporcionada por la base de datos, en la que posibles prácticas en contra de la misma resultaría en un abuso de confianza del que se habló en el punto uno. Lo anterior no es el caso de esta solución particular ya que se reconoce la fuente de información, cumpliendo con las licencias necesarias.

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

Una de las maneras más fáciles en la que mi proyecto podría caer en una falta ética en el panorama de la normas sería a través de la manipulación de datos de forma errónea a propósito por ejemplo creando un modelo que intencionalmente tuviera resultados erróneos. Otra manera podría ser utilizar los datos como si fueran mis propios datos o sin atribuir correctamente al autor. Los casos ocasionados por malicia o negligencia giran alrededor de faltas de mala fé en las que no se respeten las normas con las que se ha trabajado.

5. Liga al repo:

https://github.com/Julian7312/A00832272_PortafolioAnalisis/tree/main

Referencias bibliográficas:

1. Cortez,Paulo, Cerdeira,A., Almeida,F., Matos,T., and Reis,J.. (2009). Wine Quality. UCI Machine Learning Repository. <https://doi.org/10.24432/C56S3T>.
2. Creative Commons. (n.d.). Attribution 4.0 International (CC BY 4.0). <https://creativecommons.org/licenses/by/4.0/>