



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

Andre Sebastian Galindo Posadas A00833376

Inteligencia artificial avanzada para la ciencia de datos I (Gpo 102)

Profesor:

Alfredo Esquivel Jaramillo

Tecnológico de Monterrey

Martes 9 de Septiembre 2024

El dataset utilizado esta en el siguiente link:

<https://www.kaggle.com/datasets/dansbecker/powerlifting-database>

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

El dataset que yo utilice o que estoy analizando es sobre competencias de powerlifting, donde se ve lo que son los participantes, las categorías a las que pertenecen y además sus respectivos levantamientos en los 3 ejercicios básicos así como el lugar en el que quedaron, Ahora bien en este caso siempre antes de participar en estas competencias te hacen firmar un acuerdo en el que le cedas los datos a la competencia, esa misma información siempre se suele publicar para promocionar el evento o bien para otro tipo de análisis de este tipo de competencias, entonces estos datos ya de por sí son algo públicos, de igual manera, por lo que pude investigar en USA hay una ley que protege este tipo de información llamada **Ley de Privacidad del Consumidor de California (CCPA)**. Este marco legal regula el uso, almacenamiento y procesamiento de datos personales. Si los datos del dataset de powerlifting contienen información demasiado personal, es esencial garantizar el consentimiento explícito de los sujetos y aplicar medidas de anonimización o seudonimización para proteger su privacidad.

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.

Solo los use para entrenar un modelo de IA y entender el funcionamiento de este, realmente se podría haber hecho de cualquier cosa, pero decidí usar este dataset por gusto personal, pero realmente los datos que use, eran los máximos de los 3 levantamientos para predecir el peso de la persona con una regresión lineal, pero como tal toda la demás información como nombre, edad, categoría o lugar en el que quedó no se usó, entonces realmente no estoy usando ninguna información personal de los participantes.

- 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.**

El modelo de regresión lineal utilizado en este análisis no maneja datos personales de los participantes, ya que su propósito es predecir el peso de la persona basándose exclusivamente en sus tres levantamientos máximos en los ejercicios básicos de powerlifting (sentadilla, press de banca y peso muerto). Dado que no se utiliza información que identifique de manera directa a los individuos (como nombre, sexo, o cualquier otro dato personal), el modelo no incurre en ninguna violación de privacidad.

Además, no existen sesgos éticos significativos asociados a este modelo, dado que los datos utilizados se limitan a métricas deportivas objetivas, y no hay factores relacionados con la identidad o características personales de los participantes. Si bien el porcentaje de precisión del modelo puede variar en función de los casos individuales, este margen de variación es normal y esperado en el contexto de la predicción basada en parámetros físicos, alineándose con los estándares aceptados en la industria actualmente.

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

Se podría utilizar en el beneficio de ciertas personas al querer corromper el sistema y hacer estimaciones más altas de las que realmente fueron hechas, o incurrir en la violación de los datos personales de cada individuo que participa en esta competencia, además de todo esto se puede falsificar y a la vez vender información personal a otras personas, por lo que puede ir por allí en el caso de la malicia y la mala utilización de mi modelo de regresión lineal. Ahora en la parte de negligencia el hecho de no darle un mantenimiento adecuado o no ajustar y refinar el modelo puede llegar a causar complicaciones en un futuro, lo cual sería peligroso a la hora de hacer las predicciones.