

# **Análisis del contexto y la normatividad**

## **Normativa de datos**

Para el entregable “Uso de framework o biblioteca de aprendizaje máquina para la implementación de una solución”, se usó el dataset de “Raisin\_Dataset.xlsx”, que obtuvimos en la pagina <https://archive.ics.uci.edu/ml/datasets/Raisin+Dataset#>, el cual es un repositorio de varios datasets, que únicamente solicita que al usar los datasets de la página y crear una publicación, se citen estos para que así si alguien más quiere usar los datasets pueda hacerlo y estos ayuden a más.

## **Normativa de herramientas**

Para este entregable, únicamente se usa como herramientas para el machine learning, la librería de sklearn, la cual aunque en mi modelo, debido a lo que se está buscando predecir y los datos ingresados, no podría ser empleado para realizar algún sesgo o discriminación, pero esto no quiere decir que la herramienta no pueda ser usada y que esto se dé, un ejemplo de esto sería si en un dataset se considera el color de piel de las personas o su nacionalidad para obtener un resultado, y con esto el modelo intuya que el color de piel o nacionalidad son factores clave para la predicción, cuando realmente pueden ser solo correlación, mas no una causalidad, un ejemplo de esto sería un dataset donde se considera a la población china contra la de un país con poca población, al tener tantos registros de población china, el modelo puede tomar esto como causalidad de lo que quiere predecir, y realmente esto no afecte. Para eliminar esto, lo que se debe hacer es, antes de ingresar un dataset a tu modelo, hacer un análisis y limpieza de datos donde busquemos evitar esto.

## **Referencias**

- Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.
- About us. (2013). Scikit-Learn. <https://scikit-learn.org/stable/about.html>