

Análisis del contexto y la normatividad

Daniel Saldaña Rodríguez
A00829752

Para este reporte serán utilizados los datos del módulo de estadística.

La base de datos representa una investigación de parte de una empresa de consultoría con el objetivo de identificar los principales factores de los que depende el precio de los automóviles. Ya que la información que está siendo recopilada es pública y por lo tanto no es sensible, no hay necesidad de asegurarnos de que se mantenga segura como podría ser en otros casos. Este dataset tiene una licencia en el dominio público por lo que puede ser usado libremente, sin restricciones. (Kaggle, 2022)

En cuanto a herramientas utilizadas para el manejo de la información, se están utilizando plataformas en la nube, específicamente Google Colab y R Cloud. Estas herramientas son manejadas por empresas prestigiosas que se aseguran que los proyectos que alojan no sean comprometidos, además para acceder a estos se necesita un inicio de sesión, el cual puede ser hecho más seguro con los protocolos de verificación doble.

Sin embargo, errores de usuario pueden romper esta seguridad y poner en riesgo los datos, por ejemplo, es el caso de usar una de estas herramientas en una computadora pública y no cerrar sesión al terminar, cualquier persona podría acceder a la plataforma y tener control total. Es por esto que es tan importante no ser negligentes y darle una alta prioridad a la seguridad.

Aunque nuestro modelo está siendo entrenado con datos libres, podría llegar a ser usado de formas no éticas. Por ejemplo, el usar el modelo en datos privados que no se tiene licencia para utilizar, por lo que siempre tiene que tener un cuidado con la forma en la que utiliza un modelo de aprendizaje.

Kaggle. (2022, October 2). *Common license types for datasets*. from <https://www.kaggle.com/discussions/general/116302>