Business understanding:

En este proyecto se van a obtener datos de 3 fuentes de datos sobre distintos modelos de coches de la marca Tesla. La idea es obtener los datos mediante webscrapping y después usar esos datos para entrenar un modelo de regresión para intentar obtener el precio óptimo para un vehículo dependiendo del modelo, el color y el kilometraje.

Es importante que el modelo indique el mejor valor posible teniendo en cuenta las características para ajustar el precio lo mejor posible al precio de mercado del vehículo.

Requisitos:

Obtener los datos de necesarios de 3 fuentes distintas de datos.

El modelo tiene que obtener el precio de un vehículo dependiendo del modelo, el kilometraje y el color.

Los datos tienen que ser almacenados en una base de datos.

Restricciones:

Los datos tienen que obtenerse de fuentes abiertas de datos.

Los datos no pueden ser generados aleatoriamente.

Data understanding

Los datos que se van a utilizar para el proyecto son:

-Modelo: Modelo del vehículo tesla.

-Precio: Precio del vehículo.

-Color: Color del vehículo.

-Kilometraje: Km hechos por el vehículo. En caso de que no aparezca dicho dato se tomará 0 de referencia.

Las paginas de las que se van a obtener los datos son:

Página oficial de Tesla.

Página Milanuncios.es.