**Guion de la Presentación: Predicción del Valor de Coches de Segunda Mano en España**

**1. Introducción: La situación actual del mercado de coches de segunda mano en España**

* **Datos recientes**: Aumento de la demanda de coches de segunda mano debido a la crisis económica, la escasez de materias primas para fabricación y los altos precios de los coches nuevos. Además, sumamos la transición a un parque automovilístico más limpio, con la promoción de coches eléctricos e híbridos que aumentan el precio y que limitan el uso de coches antiguos por normativa anticontaminación.
* **Factores que influyen en los precios**: Marca, modelo, antigüedad, kilometraje, tipo de combustible, etiqueta medioambiental, potencia, etc.
* **Problema principal**: La variabilidad en los precios hace que tanto compradores como vendedores tengan dificultades para establecer precios justos.
* **Objetivo del estudio**: Crear un modelo de Machine Learning que prediga el valor de un coche de segunda mano con precisión, ayudando a mejorar la transparencia del mercado.

**2. Enfoque del problema y planteamiento del modelo**

* **Recogida de datos**: se ha utilizado un dataset de DataMarket que recoge la información de las principales páginas online de venta de vehículos de segunda mano en España.
* **Características utilizadas (features)**: Variables más relevantes y por qué se han seleccionado.
* **Limpieza y preprocesamiento de datos**: Tratamiento de valores nulos, codificación de variables categóricas, normalización.

**3. Modelado y entrenamiento del modelo**

* **Modelos probados**: Comparación entre regresión lineal, Random Forest, LightGBM, XGBoost y CatBoost.
* **Evaluación del modelo**: Métricas utilizadas (RMSE, R²) y rendimiento del mejor modelo.
* **Optimización y ajuste de hiperparámetros**: Técnicas como Grid Search o Random Search.

**4. Resultados y conclusiones**

* **Precisión del modelo**: Cómo se comporta en la predicción de precios.
* **Casos de uso**: Cómo podría beneficiar a concesionarios, vendedores y compradores.
* **Limitaciones**: Aspectos a mejorar y posibles sesgos en los datos.
* **Futuras mejoras**: Inclusión de más datos, modelos más complejos, impacto de la oferta y la demanda. Dividir dataset en tipos de vehículo (utilitarios de gama baja-media, coches de lujo, deportivos, etc.) para crear diferentes modelos según el tipo, evitando así outliers.

**5. Demostración (opcional)**

* Presentación de un ejemplo real utilizando el modelo para predecir el precio de un coche con ciertas características.

**Introducción: La situación actual del mercado de coches de segunda mano en España**

El mercado de coches de segunda mano en España ha experimentado un crecimiento significativo en los últimos años, impulsado por diversos factores económicos y sociales. La crisis de los semiconductores, que ha afectado la producción de vehículos nuevos, ha generado un incremento en la demanda de vehículos usados, haciendo que sus precios se disparen. Según datos recientes, en 2023 el precio medio de los coches de segunda mano en España aumentó en torno a un 10% con respecto al año anterior, alcanzando cifras récord.

Este encarecimiento no solo ha complicado el acceso a vehículos asequibles para los consumidores, sino que también ha generado incertidumbre en los vendedores. Determinar un precio justo para un coche usado es una tarea compleja, ya que depende de múltiples factores como la marca, el modelo, la antigüedad, el kilometraje, el tipo de combustible, la etiqueta medioambiental, la potencia del vehículo, etc.

En este contexto, surge la necesidad de herramientas que ayuden a predecir el valor de los coches de segunda mano de manera precisa y objetiva. La aplicación de técnicas de Machine Learning permite analizar grandes volúmenes de datos y detectar patrones que pueden mejorar la estimación del precio de un vehículo, beneficiando tanto a compradores como a vendedores.

Con este objetivo, se ha desarrollado un modelo de Machine Learning que predice el precio de un coche de segunda mano en función de sus características. En esta presentación, explicaré cómo se ha planteado el problema, qué enfoques se han utilizado y cuáles han sido los resultados obtenidos.