

Descripción del proyecto

¡Felicidades! Has completado otro curso de la plataforma de entrenamiento. Ahora es el momento perfecto para probar tus habilidades y resolver un nuevo problema de machine learning. Al realizar este proyecto, trabajarás por tu cuenta.

Cuando termines, envía tu trabajo al revisor del proyecto. Recibirás feedback dentro de las siguientes 24 horas. Después de esto, harás los cambios necesarios a tu trabajo y lo enviarás para una segunda revisión.

Por lo general, este proceso se repetirá varias veces hasta que recibas el visto bueno del revisor y se aprueben todas las correcciones. Todo eso es parte del trabajo.

Tu proyecto se considerará completado una vez que el revisor del proyecto lo apruebe.

Descripción del proyecto

Rusty Bargain es un servicio de venta de coches de segunda mano que está desarrollando una app para atraer a nuevos clientes. Gracias a esa app, puedes averiguar rápidamente el valor de mercado de tu coche. Tienes acceso al historial, especificaciones técnicas, versiones de equipamiento y precios. Tienes que crear un modelo que determine el valor de mercado.

A Rusty Bargain le interesa:

- la calidad de la predicción
- la velocidad de la predicción el tiempo requerido para el entrenamiento
- el tiempo requerido para el entrenamiento

Instrucciones del proyecto

1. Descarga y examina los datos.
2. Entrena diferentes modelos con varios hiperparámetros.
3. Analiza la velocidad y la calidad de los modelos.

Observaciones:

- Utiliza la métrica *RECM* para evaluar los modelos.

- Aprende por tu propia cuenta sobre la librería LightGBM y sus herramientas para crear modelos de potenciación del gradiente (gradient boosting).
- Puedes utilizar un comando especial para descubrir el tiempo de ejecución del código de una celda en Jupyter Notebook. Encuentra ese comando.
- Como el entrenamiento de un modelo de gradient boosting puede llevar mucho tiempo, límtate a cambiar unos pocos parámetros del modelo.
- Si Jupyter Notebook deja de funcionar, elimina las variables excesivas utilizando el operador `del`:

```
del features_train
```

Descripción de los datos

El dataset está almacenado en el archivo `/datasets/car_data.csv` . [Descargar dataset](#).

Características

- *DateCrawled*: fecha en la que se descargó el perfil de la base de datos
- *VehicleType*: tipo de carrocería del vehículo
- *RegistrationYear*: año de matriculación del vehículo
- *Gearbox*: tipo de caja de cambios
- *Power*: potencia (CV)
- *Model*: modelo del vehículo
- *Mileage*: kilometraje (medido en km de acuerdo con las especificidades regionales del conjunto de datos)
- *RegistrationMonth*: mes de matriculación del vehículo
- *FuelType*: tipo de combustible
- *Brand*: marca del vehículo
- *NotRepaired*: vehículo con o sin reparación
- *DateCreated*: fecha de creación del perfil
- *NumberOfPictures*: número de fotos del vehículo

- `PostalCode`: código postal del propietario del perfil (usuario)
- `LastSeen` — fecha de la última vez que el usuario estuvo activo

Objetivo

Price - precio (en euros)

Evaluación del proyecto

Hemos definido los criterios de evaluación para el proyecto. Léelos con atención antes de pasar al ejercicio.

Esto es en lo que se fijarán los revisores al examinar tu proyecto:

- ¿Seguiste todos los pasos de las instrucciones?
- ¿Cómo preparaste los datos?
- ¿Qué modelos e hiperparámetros consideraste?
- ¿Conseguiste evitar la duplicación del código?
- ¿Cuáles son tus hallazgos?
- ¿Mantuviste la estructura del proyecto?
- ¿Mantuviste el código ordenado?

Ya tienes tus hojas informativas y los resúmenes de los capítulos, por lo que todo está listo para continuar con el proyecto.

¡Buena suerte!