

# 二手车价格预测

## 任务和主题

预测二手车的交易价格。

## 数据说明

该数据来自某交易平台的二手车交易记录，包含 31 列变量信息，其中 15 列为匿名变量。已对 name、model、brand 和 regionCode 等信息进行脱敏。

### 数据字段表

Field	Description
SaleID	交易ID，唯一编码
name	汽车交易名称，已脱敏
regDate	汽车注册日期，例如20160101，2016年01月01日
model	车型编码，已脱敏
brand	汽车品牌，已脱敏
bodyType	车身类型：豪华轿车：0，微型车：1，厢型车：2，大巴车：3，敞篷车：4，双门汽车：5，商务车：6，搅拌车：7
fuelType	燃油类型：汽油：0，柴油：1，液化石油气：2，天然气：3，混合动力：4，其他：5，电动：6
gearbox	变速箱：手动：0，自动：1
power	发动机功率：范围 [ 0, 600 ]
kilometer	汽车已行驶公里，单位万km
notRepairedDamage	汽车有尚未修复的损坏：是：0，否：1
regionCode	地区编码，已脱敏
seller	销售方：个体：0，非个体：1
offerType	报价类型：提供：0，请求：1
creatDate	汽车上线时间，即开始售卖时间
price	二手车交易价格（预测目标）
v系列特征	匿名特征，包含v0-14在内15个匿名特征

- 训练集：used\_car\_train.csv (used\_car\_train.zip)，15 万条。
- 测试集 A：used\_car\_testA.csv (used\_car\_testA.zip)，5 万条，不包含 price 字段。
- 测试集 B：used\_car\_testB.csv (used\_car\_testB.zip)，5 万条，不包含 price 字段。

## 结果提交

预测测试集 AB 中二手车价格，补充 used\_car\_submitA.csv 以及 used\_car\_submitB.csv 中 price 列。格式如下：

SaleID	price
150000	687
150001	1250