

二手车估计问题分析

1. 问题分析：

随着机动车数量增长，二手车流通愈加频繁，流通手段包括二手车收车，拍卖，二手零售，二手车置换等环节。这种特殊的商品因为其大量的特征，难以准确估算价值和设定。品牌，车系，动力，基础配置，车况行驶里程，车身受损，维修情况，甚至新车价格都会对二手车价格带来影响。目前国家并没有出台完整的评价标准，此时构建一个公认的准确的，完整的价格估算模型尤为重要。

2. 可能会在问题三出现的思考角度：

在典型二手车零售交易中，也存在一系列问题如：获取用户线索，线下实体门店对外展销和售卖，收购价格与出售价格，收购时间与出售时间，滞销解决方案等一系列角度。

3. 问题一的分析过程

问题一要求基于给定的二手车交易样本数据作为训练，验证，测试样本。构建二手车零售交易价格预测模型。首先要求对数据的处理，脱敏过程，很传统的数据分析流程；

- 特征归一化
- 类别性特征
- 高维组合特征处理

处理数据后，进行最为重要的特征工程（因为属性比较多可以建议尝试一下PCA降维处理）或者，最支持的一种方式就是搭建神经网络将数据标准化，样本大概三万多条可以尝试搭建多个隐藏层的网络，网络最后一层不设置激活（标量回归是预测单一连续值回归经典设置），可以预测任意范围的值。同时编译网络使用MSE损失函数，监控的指标为MAE。评估网络后使用K折验证，计算验证分数平均值，绘制验证分数曲线。训练最终模型。对比与实际价格偏差度。

4. 问题二的分析过程

第二问是很传统的影响因素分析问题，题目给出了可能影响销售成交周期的关键因素，分析那些方法可以加快车辆销售速度。并证明其适用条件和预期效果。

建议第一问使用深度学习方法进行预测，第二问进行较为深入的特征工程，分析每个特征对成交周期的影响因素，可以使用SPSS等方法做关联性分析，进一步考虑多个特征结合的条件下关联性的变化情况问题。这一段比较靠文笔同时也需要一些尝试练习。

5. 问题三

问题三类似与拓展思路，可以结合问题一，问题二的一些结果分析，问题二可能的关联性分析但是无法给出实际问题的可能进行探讨，类似给公司的一封信，做总结陈述，以及未来建议。

6. 结论

总的来说，MothorCup提出的两个题目都是比较商业性质的题目。个人建议是第一问使用深度学习构建预测模型，第二问使用关联分析，特征工程进行讨论，构建多项式预测模型，对比实际成交价格证明模型的可靠性。