**数据建设总结**

**迫切用户区分**

**需求：**业务反馈高价楼盘中存在很多非高消费人群的标签，使得报告结果与实际情况相差较大，需调整

**过程：**取部分高价楼盘的用户标签进行校验，发现很多低消费标签用户有低频次点击高价盘的行为，故做了算法优化，保留原始真实点击数据的前提下，根据楼盘均价所处城市均值的水平，划分为高、中、低三挡楼盘，再结合用户居住小区价格，相同楼盘点击频次，线下消费场景划分消费能力高、中、低三挡与楼盘对应，匹配核心迫切用户人群。

**结果：**原迫切用户算法是基于用户点击形同楼盘频次来计算，本次算法修后提升了迫切用户契合度。

**标签优化**

**需求：**标签质量不高

**过程：**核查标签判定算法，发现部分标签供应商的算法存在漏洞，例如驾车出行标签判断逻辑，发现判断用户有车的逻辑主要取决于个别APP是否安装，结果判定过于片面，故做了判断算法优化，根据标签属性，添加居住地与工作地的距离+用户年龄+消费能力+线下消费场景出行频次等标签综合判断，不在局限个别APP是否安装。

**结果：**通过这个问题发现其他标签也存在这类问题，协同供应商一起调整了多个有问题的标签，提升了标签准确率

**竞对算法优化**

**需求：**竞对算法优化

**过程：**与业务沟通竞对核心竞争力，不在局限户型、价格、区域的属性类比，增加同一用户还看过其他哪些类似楼盘的浏览对比，并以多维散点图的形式做呈现，使用机器学习算法处理庞大的计算量得到最终计算结果

**结果：**竞对对比呈现更为清晰直观，对业务判断起到很大的支持