



中國人民大學
RENMIN UNIVERSITY OF CHINA

房间预测项目展示

汇报人：谢丽媛

2025年4月

"建筑面积"和"套内面积"预处理方案

❑ 衍生特征:

得房率: 套内面积/建筑面积 (核心指标)

公摊面积: 建筑面积-套内面积

面积等级: 按市场标准分段 (50/90/144/200m²为界)

是否豪宅: 144m²为分界线 (中国普通/非普通住宅标准)

❑ 缺失值处理:

使用中位数得房率估算缺失的套内面积

重新计算得房率确保一致性

❑ 异常值处理:

确保套内面积 ≤ 建筑面积

得房率合理范围约束 (后续可添加)

关键处理逻辑

方向标准化:

模型亮点

训练集"房屋户型"预处理方案

▣ 异常值处理:

卧室数超过10的设为10 (应对"16室"等异常)

卫生间数不超过卧室数+3 (应对"14卫"等异常)

▣ 衍生特征:

总房间数: 卧室+客厅

卧室卫生间比: 反映居住舒适度

是否标准户型: 标记功能齐全的户型

是否豪宅户型: 卧室 ≥ 4 的户型

训练集"房屋朝向"预处理方案

▣ 特征提取:

主朝向: 按价值优先级选择 (南>东南>东>...)

优质标志: 南、东南、东朝向为优质

▣ 业务逻辑编码:

中国购房者偏好: 南>东南>东>西南>北>西>东北>西北

南北通透户型有显著溢价

