

从机器学习的角度看如何家装界的Uber

Agenda

- “共享经济”创业公司们在如何使用机器学习
- 设计家的应用场景
- 如何应用机器学习

“共享经济”创业公司们
在如何使用机器学习



Find unique homes, experiences, and local guides for your trip.

Challenge

- 定价艰难
 - 如何让不专业的房东决定价格？
 - 和专业酒店模式相竞争
- 房东大量驳回预约申请，严重影响体验

Solution - 智能定价



Price Tips: Airbnb为host设计的房屋动态定价工具，向房东展示房屋在某一价格下被租出去的可能性。房东利用这个功能灵活地根据市场供需关系的变化来制定价格。

Solution - 智能定价

通过机器学习算法根据下面的维度来定价

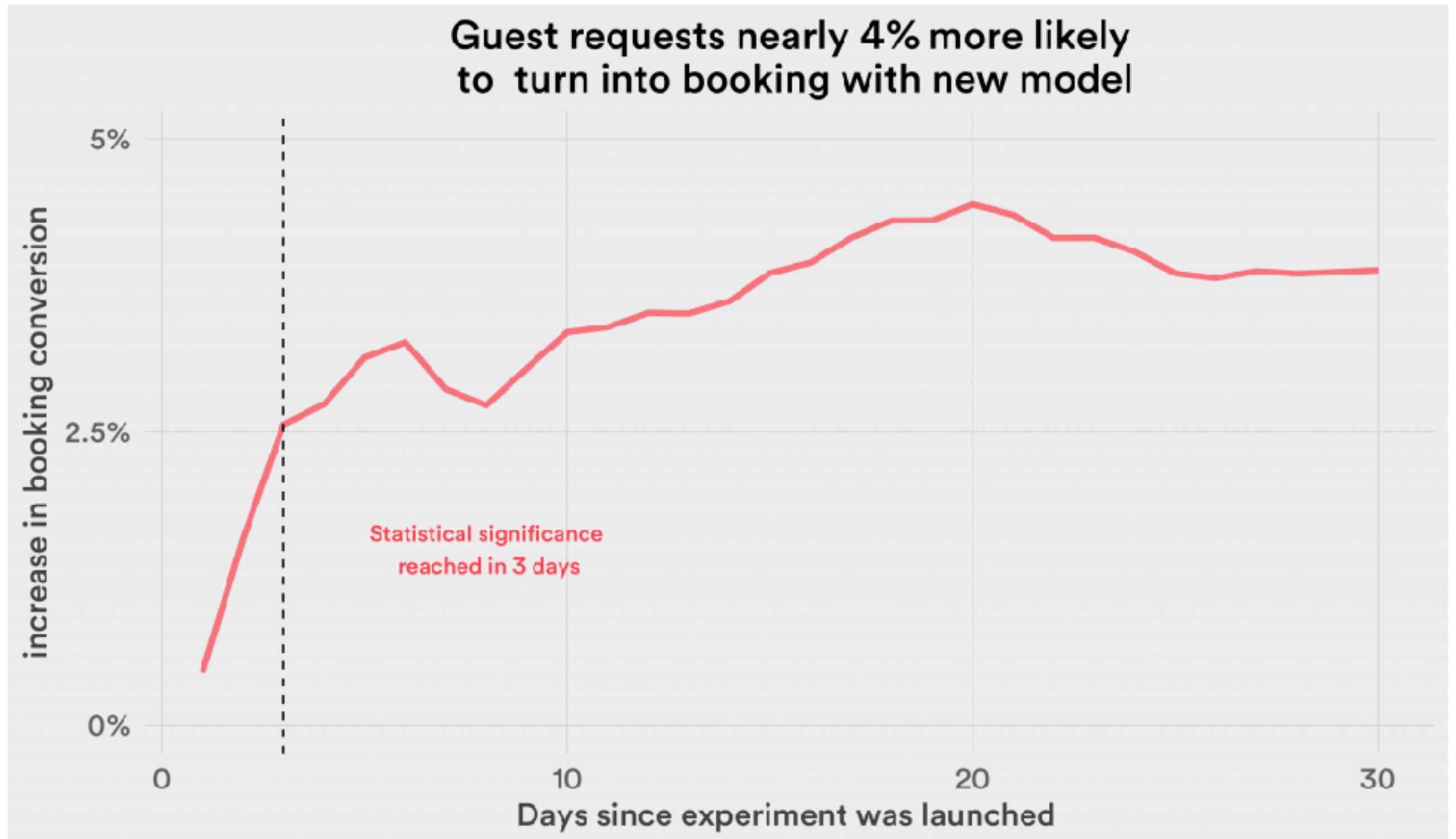
- 封面图片
- 所在小区
- 节假日
- 大型活动
-

机器学习算法的定价比人工更加专业



封面图片

Solution - 自动撮合考虑到房东偏好



在Airbnb上线了房东偏好模型后的短短几天，撮合成功率就提高了4个百分点

Tech

- Airbnb 开源的机器学习库: <https://github.com/airbnb/aerosolve>

The Uber logo, consisting of the word "UBER" in white, bold, sans-serif capital letters, centered within a solid black square.

UBER

Make transportation as reliable as running water, everywhere, for everyone.

Challenges

- 乘客等待时间长，出行体验差
 - 繁琐的议价过程
- 怎样让司机接受Uber的霸王条款
 - 替代方案似乎对司机更友好

Solution

- 全自动议价 — 极大提升乘客出行体验
- 智能分单 — 解决司机乘客匹配度问题，司机收入分配

Tech

- Machine Learning as a Service
- Boosted Trees
- Linear Learners
- Neural Networks
- Hadoop + Spark + MLlib (可扩展机器学习库).
- 一个集中化的数据仓库



滴滴出行

滴滴研究院

ETA

基于滴滴出行的海量实时出行数据，设计出全新的时间预估算法，从原理上克服了传统算法的缺陷，大幅度提升了时间预估的准确率。

智能分单

在司机和乘客的历史数据中学习接单概率模型，提高司机和乘客的匹配度，利用运力的规模效应实时地从全局上最优化总体交通运输效率和乘客出行体验。

拼车

实现高效的实时拼车订单的匹配算法，共享运力资源以及降低乘客出行成本。在不断提高拼成模型准确率的同时，优化拼车体验，提升服务效率。

供需预测

基于海量实时出行数据，以数十亿订单数据和数百万司机位置信息为基础，预测任意时间段各个区域的订单需求和运力分布状况，提供最优的未来出行方案。

运力调度

基于供需预测结果，大规模有序调动全城所有可用运力，实现资源最优化分配，力求解决正在发生的以及潜在供需失衡的状况，提升平台效率的同时最大化利用交通运力，缓解城市拥堵。

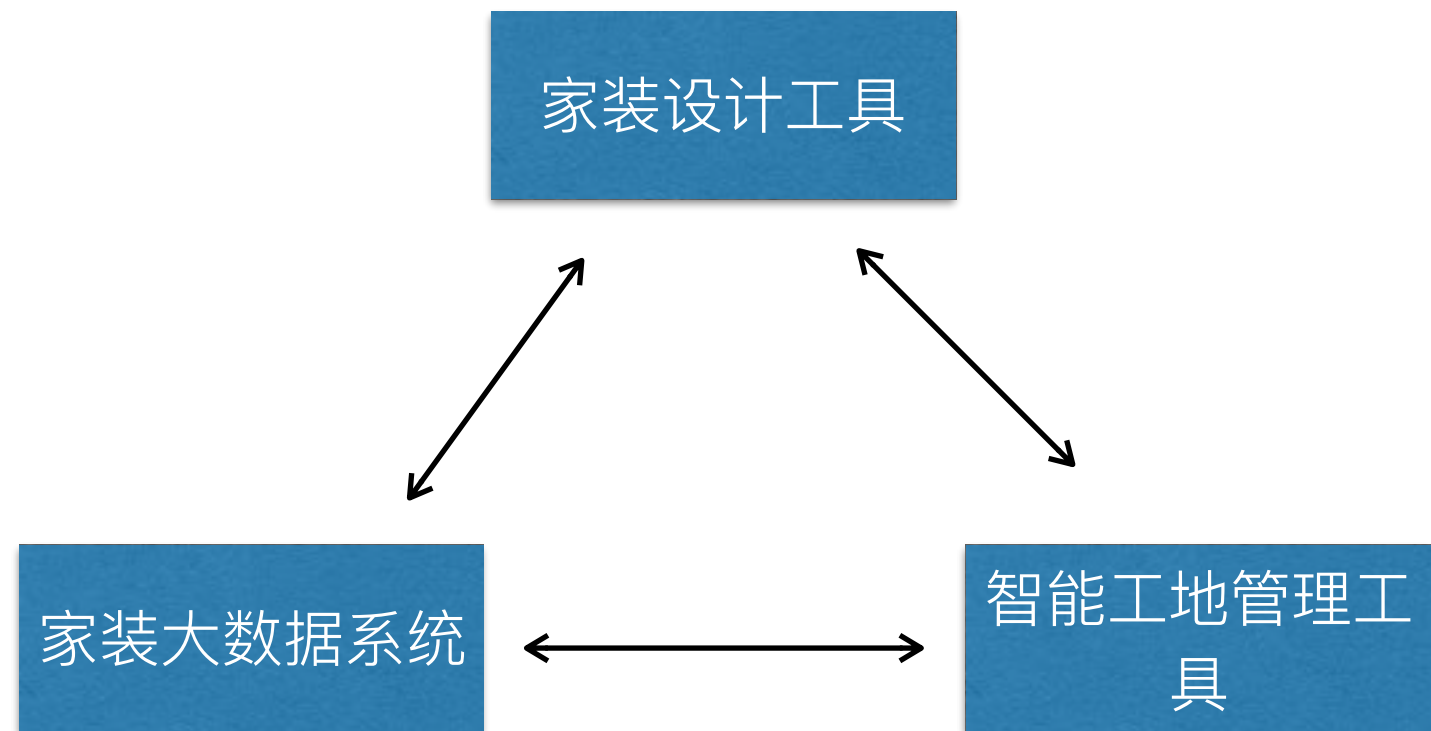
热力图

基于对历史数据的统计并结合实时订单数据，给出当前全城范围内订单密集区域的分布，给司机提供有价值的听单位置参考，提高听单概率并减少司机空驶时间。

土八兔

当前系统构成

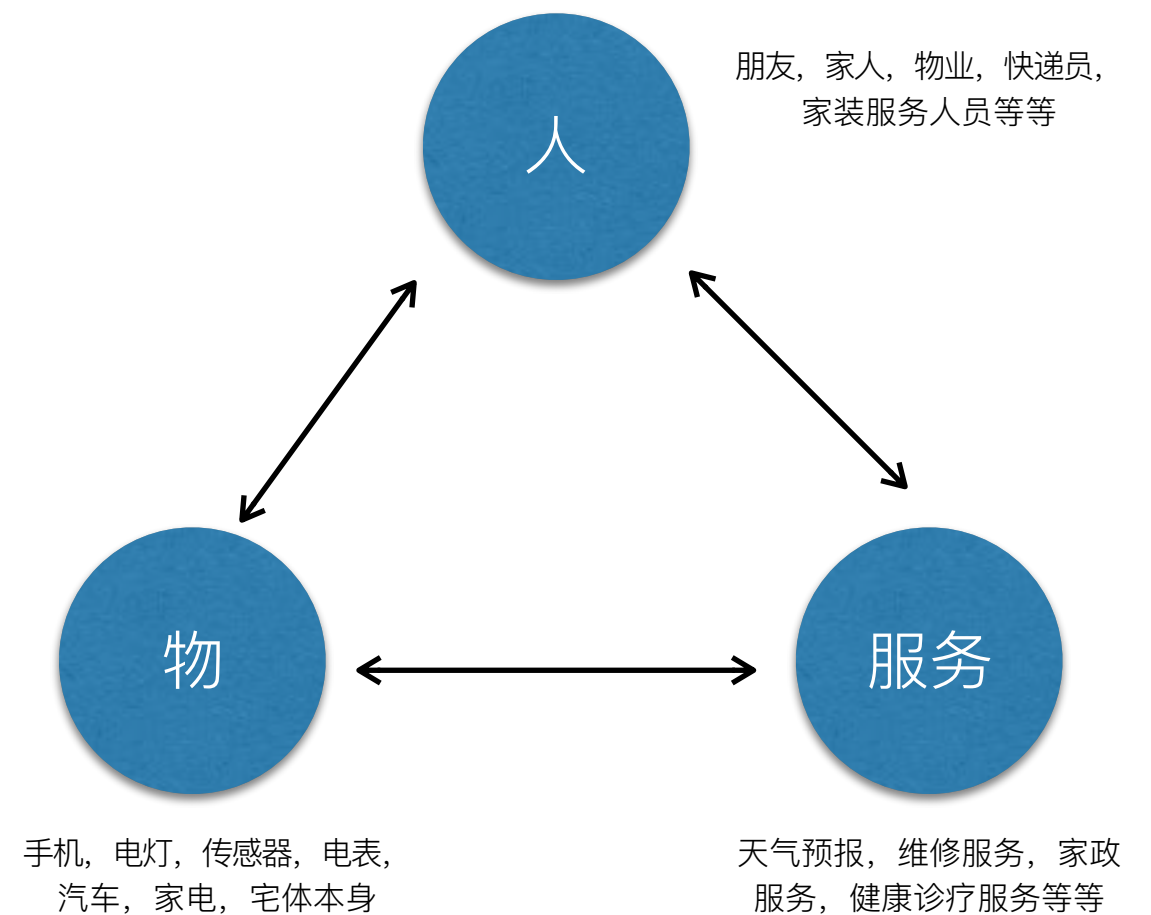
- 定位：互联网家装企业，提供一站式家装服务
- 数据：积累了1400万家装用户需求数据



土八兔

智能家居平台

- 未来想要做智能家居平台：成立了智能家居事业部。围绕智能家居诞生全新服务与行业，打造智能家居平台



畅想机器学习在设计家 的应用场景

Challenges

- 解决大规模定制化问题
 - 面对成千上万的设计师和消费者，如何将它们精确匹配
 - 如何满足消费者的个性化需求
- 设计施工标准化
 - 设计师报价没经验
 - 施工周期没有保障

Solution

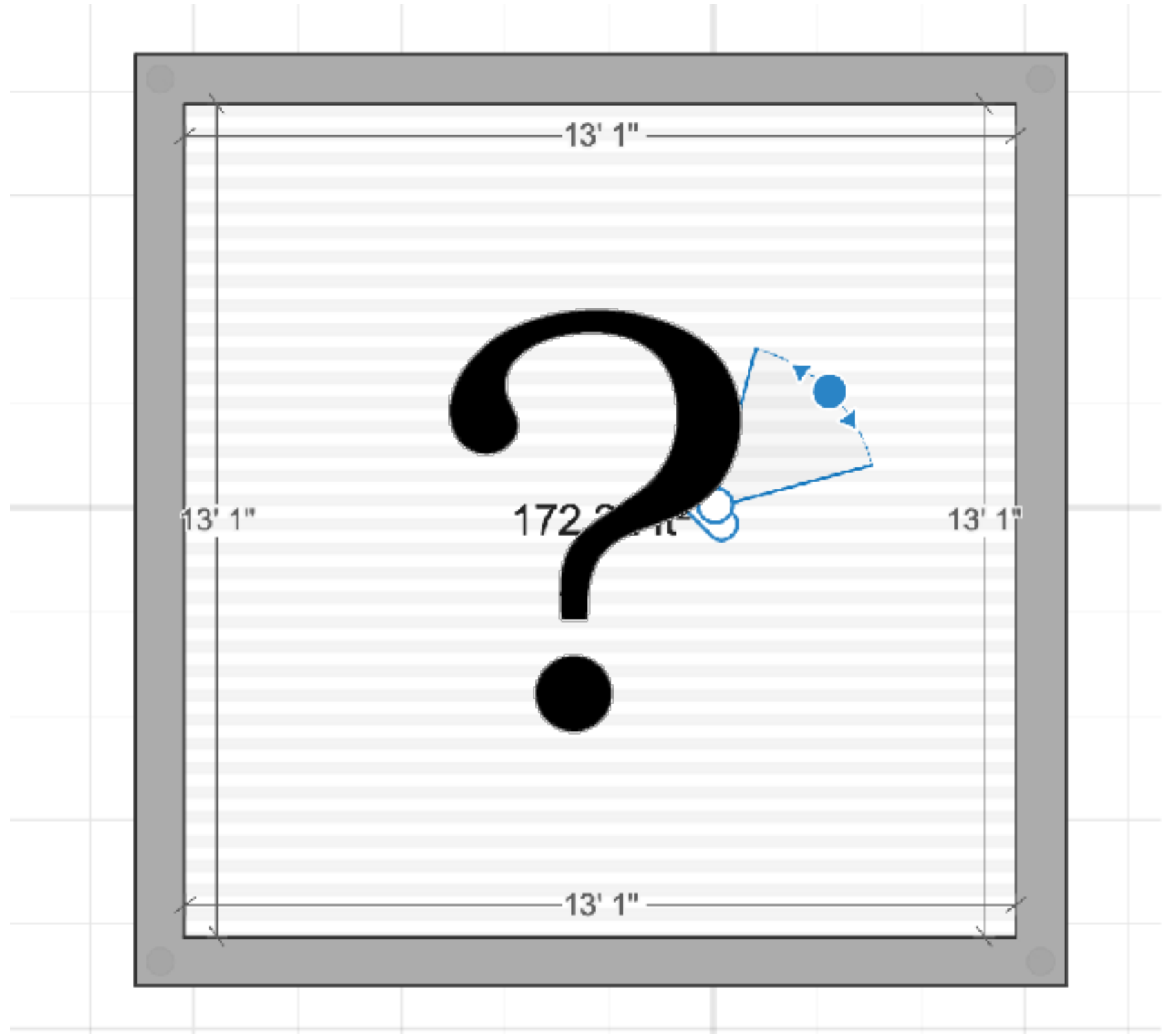
- 解决大规模定制化问题
- Solution - 智能化的3D设计工具：通过对大数据的挖掘，根据用户喜好，房型，预算等因素智能预测适合的家具、建材。增加厂家商品曝光率，提升厂家入住积极性。
- Solution - 智能派单系统（类似于uber）：通过建立消费者模型和设计师模型将两者迅速匹配。最大化双方利益，增加设计师和消费者粘性。增加设计家平台的话语权。为设计家创造上帝视角。

Solution

- 施工周期没有保障
 - Solution - 智能周期预测：类似于外卖app中的“预测抵达时间”。基于历史数据和施工需求智能预测所需时间，为用户创造合理的心理预期。
- 设计师报价没经验
 - 智能定价系统 — Solution
 - 设计师作为个体户没有专业的定价能力
 - 协助设计师做出有竞争力的定价，增加赢单几率

我们有可能像替代司机一样替代设计师吗？

- 快速导入
- 户型图智能识别
-



如何应用机器学习

Google - Rules of Machine Learning:
Best Practices for ML Engineering

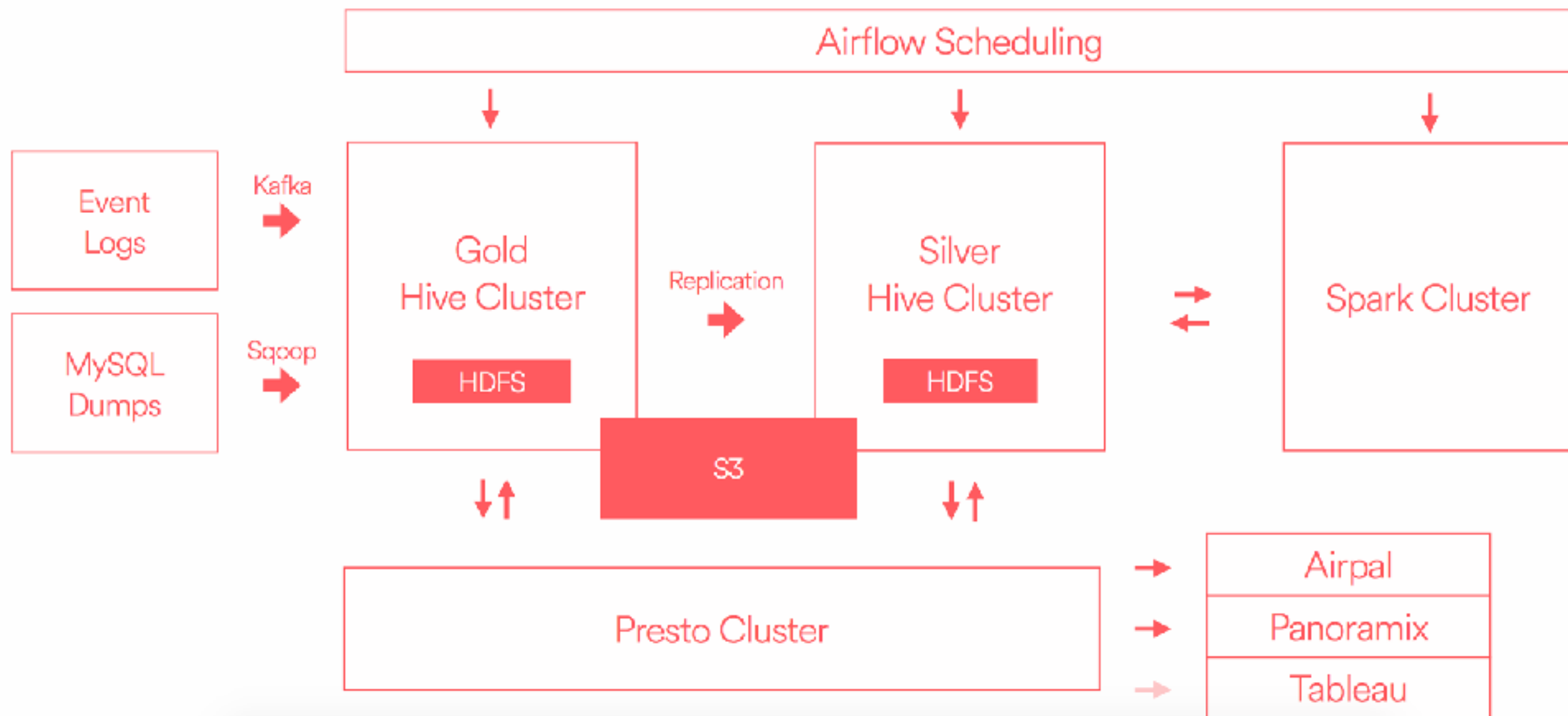
做机器学习之前

1. 对发布一个不含 ML 技术的产品，不要有顾虑
 - 如果机器学习对于你的产品非必需。在获得数据前不急着想使用它
2. 首先要设计和贯彻指标
 - 在定义ML系统要做什么之前，先尽可能多追踪你的系统
3. 在简单机器学习算法和复杂启发算法之间，选择前者

机器学习第一阶 —— 打造第一个pipeline

1. 第一个模型要简单，把基础设施弄好，决定
 - 如何获取学习算法样例
 - 定义系统的“好”与“坏”
 - 如何将模型整合进应用。实时模型还是线下模型
2. 测试基础设施要与 ML 测试分开
 - 确保基础设施可测试
3. 复制流水线时当心数据遗落

AIRBNB DATA INFRA



Q&A