



携程技术中心



IT大咖说
知识分享平台

携程技术沙龙

优化模型不得不想的几个问题 — 理论与案例

分享人（胡淦）



胡淏

- 数学/经济学/统计学 (ML)
- 美团 / 支付宝 / Ctrip / Quaera
- 算法体系、平台
- Columbia University grad

目录

CONTENTS

- 1 如何优化一个模型
- 2 模型类项目方法论
- 3 复杂业务的模型体系

桔子精选商务大床房 查看详情	标准价 预订满意度 100%	大床	无早	无线免费	2人	免费取消	¥839	预订 到店付 房量紧张
促销大床房 查看详情	标准价 预订满意度 100%	大床	无早	免费	2人	免费取消	¥639	订完 到店付
促销双床房 查看详情	标准价 预订满意度 100%	双床	无早	无线免费	2人	免费取消	¥679	订完 到店付
双床房 查看详情	标准价 预订满意度 100%	双床	无早	无线免费	2人	免费取消	¥729	订完 到店付
大床房 查看详情	标准价 预订满意度 100%	大床	无早	无线免费	2人	免费取消	¥689	订完 到店付
桔子精选豪华大床房 查看详情	标准价 预订满意度 100%	大床	无早	无线免费	2人	免费取消	¥829	订完 到店付

驾车距离28.8公里 (约78分钟)
首都国际机场
驾车距离23.6公里 (约29分钟)

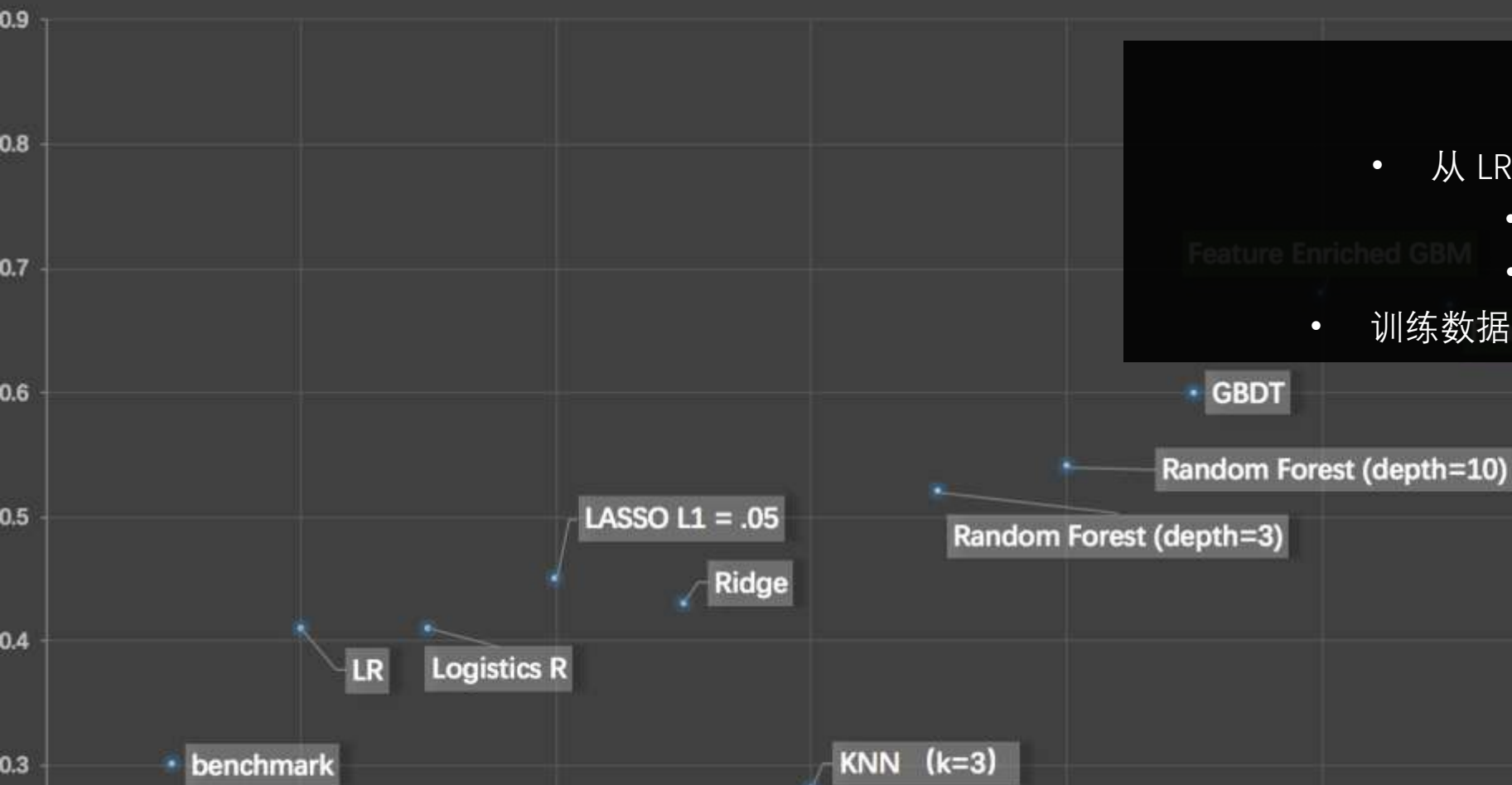
背景：

- 业务问题：
 - 10%房型不可订，
 - 数据分析发现，30%不可订房型实际可订
- 痛点：无法批量、准确识别可订房型
- 解决方案：通过机器学习周期性搜索可订房型，自动开房

我要提问

常见问题

满房开房准确率 @recall = .2



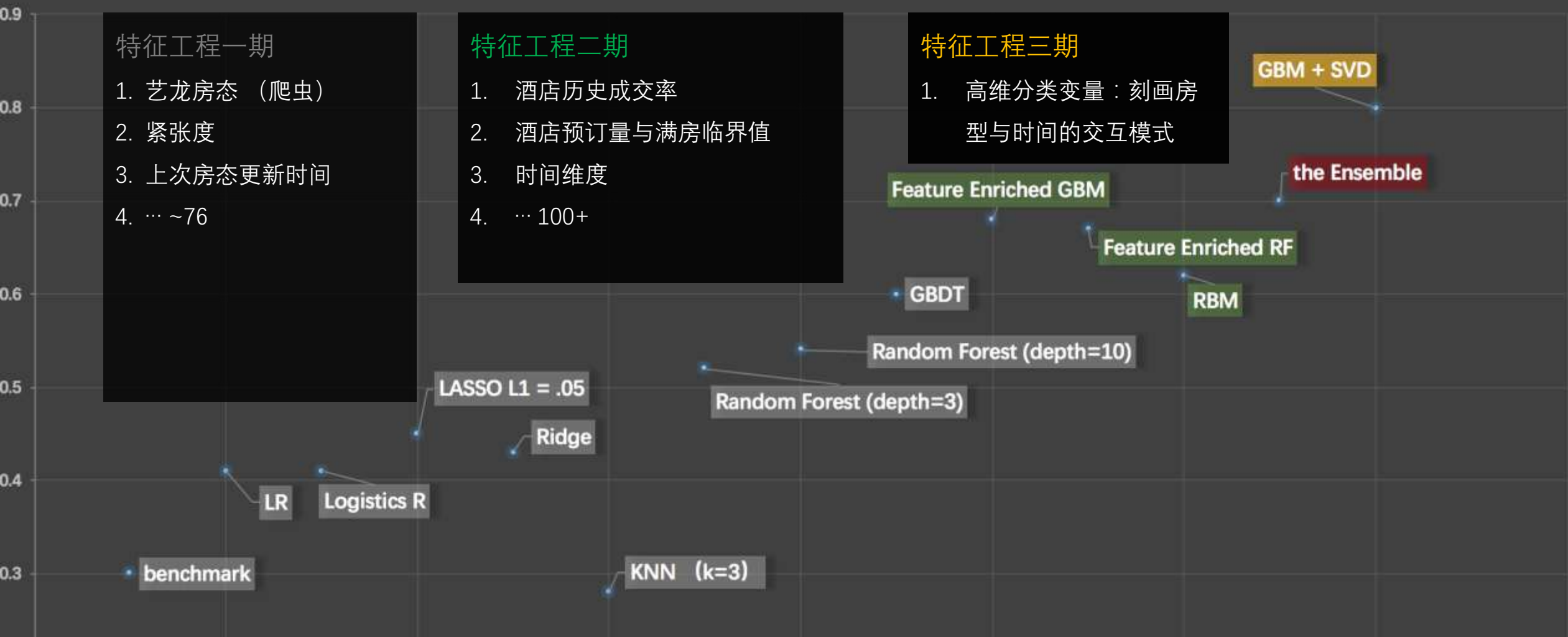
模型：

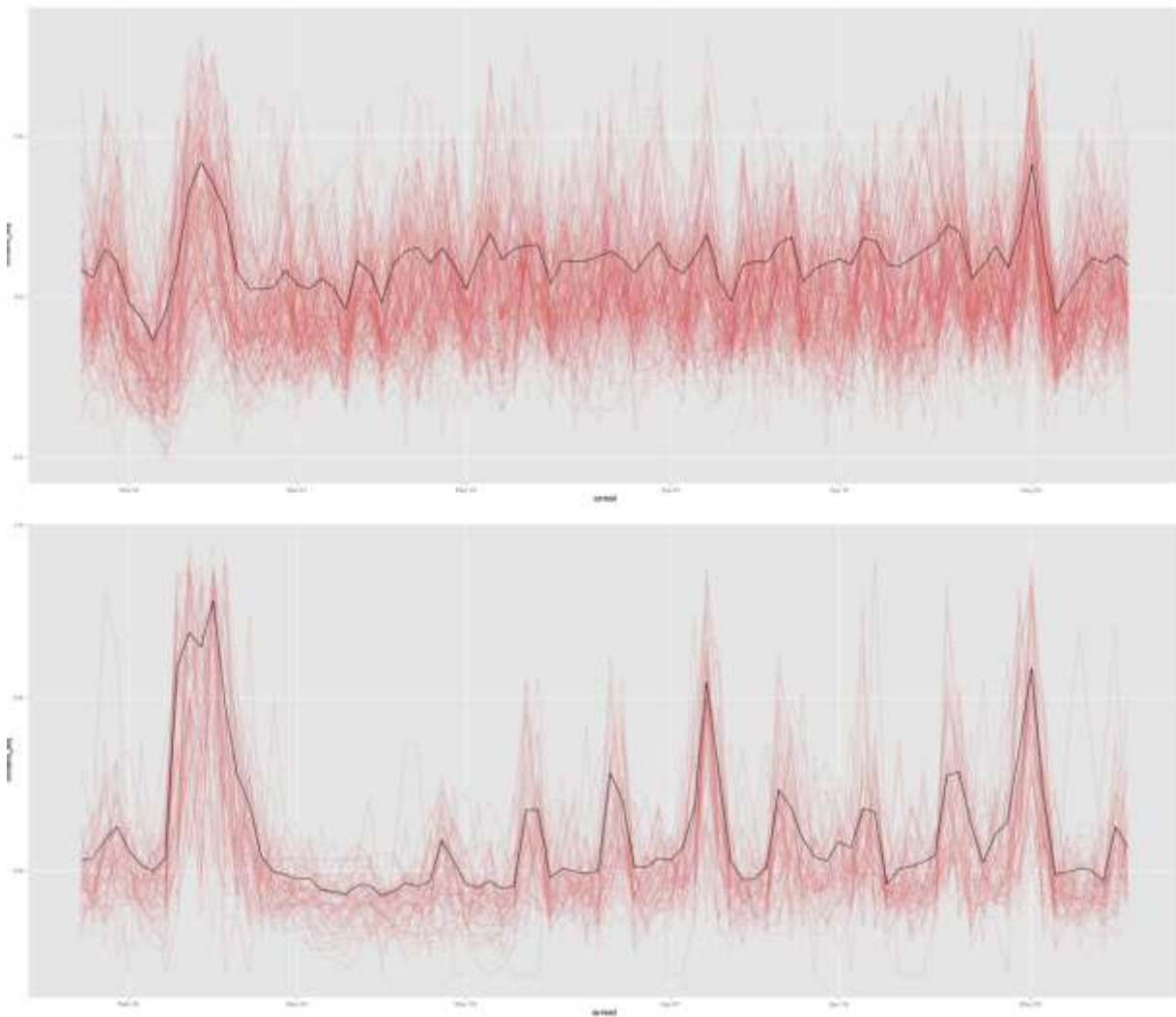
- 从 LR 到 GBDT, 从 40% 到 60%
 - RF < GBDT & xgboost
 - LASSO > LR & glmnet
- 训练数据的质量和代表性是大前提

Advice :

1. 技术选型总会遇到天花板
2. 把复杂模型和系统调参数的工作放在项目最后做

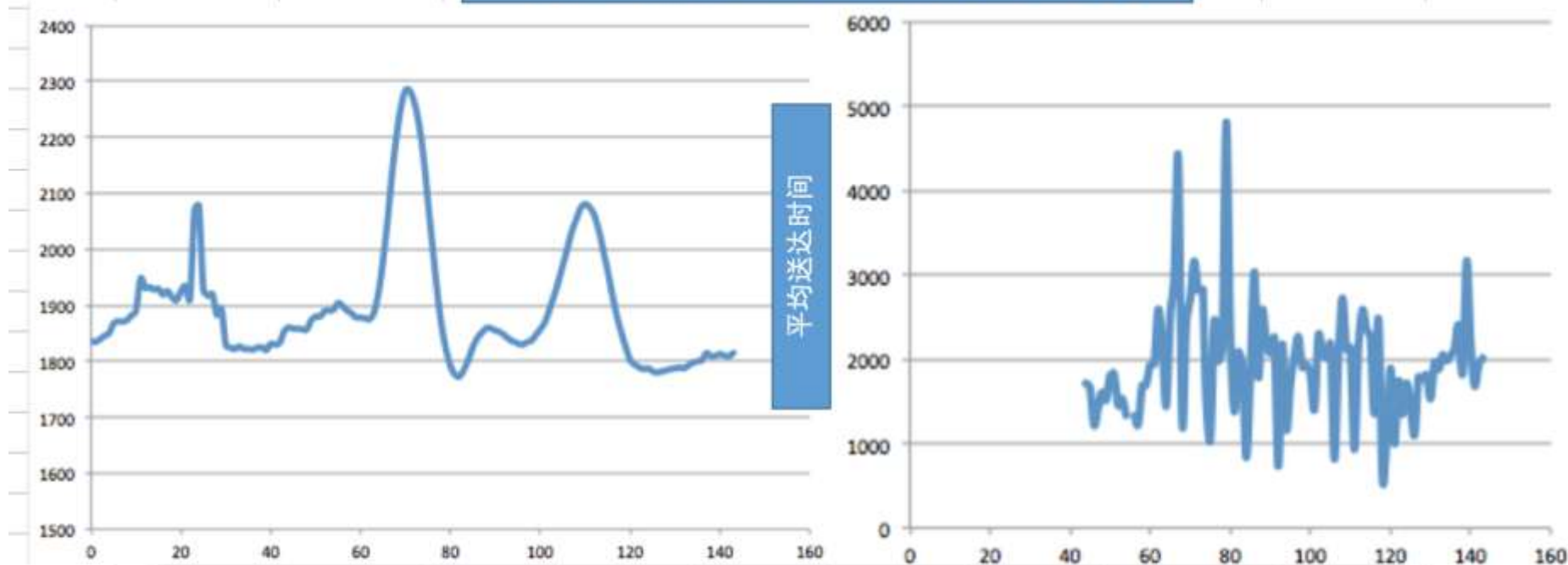
满房开房准确率 @recall = .2





1. 关键问题：一个房型当日、未来的满房率是什么样子的？
2. 业务痛点：不同房型、不同日期的满房概率是不同的；节假日预测不准确。
3. 方法
 - a) 城市-房型结构化SVD (SoftImpute)
 - a) Netflix: Rate = User + Movie + user:Movie
 - b) Ctrip: $P = \text{Room} + \text{Date} + \text{Room:Date}$
 - c) 降维: $U_{md} D_{dd} V_{nd}^T \rightarrow U_{m3} D_{33} V_{n3}^T$
 - b) 冷启动

商家到时段平均送达时间特征基于SVD的优化效果

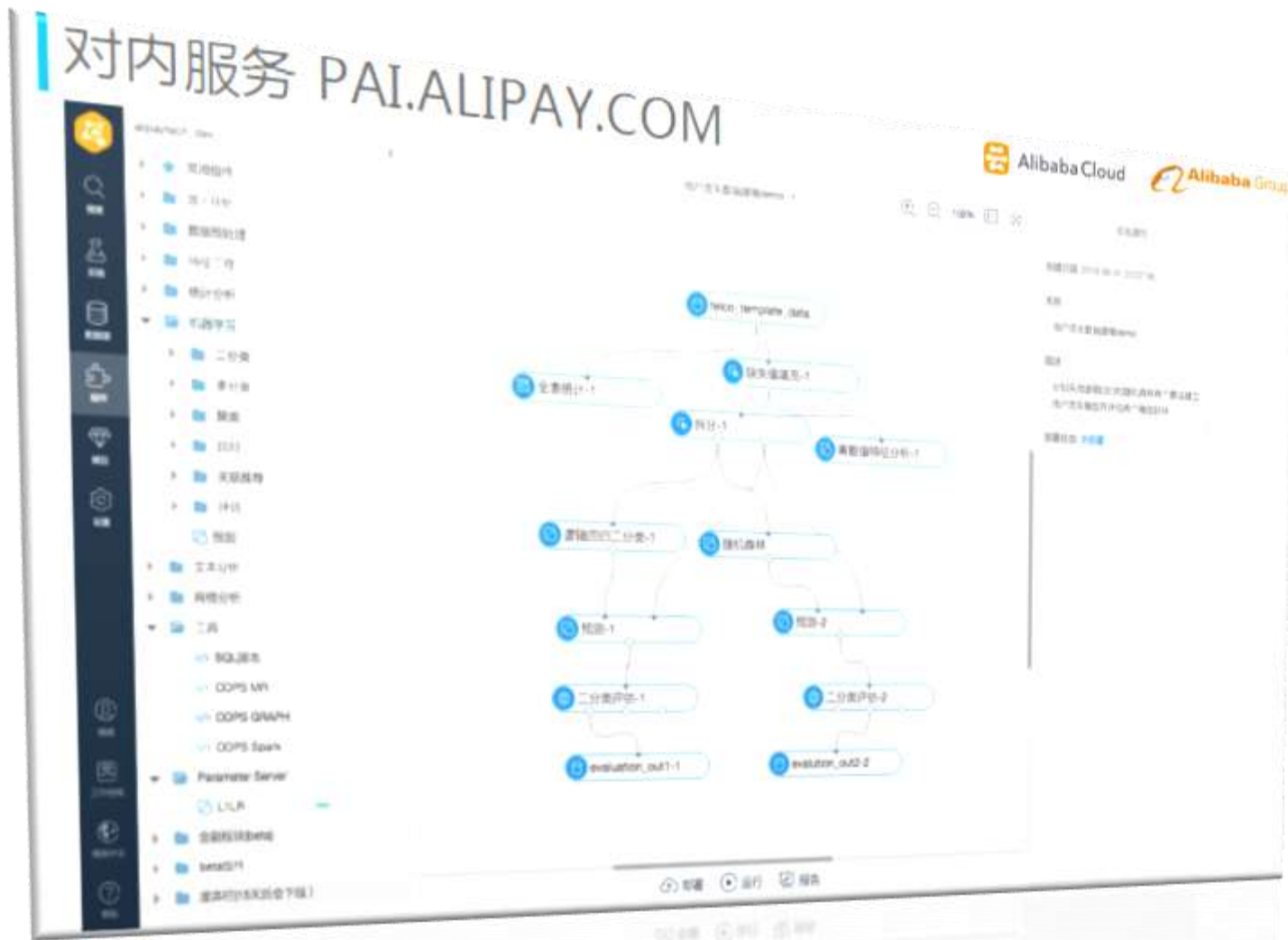


	15分钟置信度	10分钟置信度	5分钟置信度	8分钟业务准时率	15分钟业务准时率	平均误差(分钟)	有偏性(分钟)
线上现有模型	74.47%	53.59%	26.62%	50.62%	76.98%	11.10	+7.3
改进商家平均送达时长后训练的模型	85.04%	67.28%	37.44%	65.43%	86.49%	8.72	+5.0
	+10.57%	+13.69%	+10.82%	+14.81%	9.51%	-2.38	



特征工程这件事儿，要做到极致。

1. 时刻思考新特征拓展、旧特征细化的工作
 2. 业界可以做出更优秀的模型：
 - 数据不是限定的
 - 运营那边有很多信息量
 - 把业务理解转化为特征建模的过程很重要
- **RD在做特征时通常的问题**
 - 缺思路 (变量体系)
 - 缺方法 (业务>特征>模型)
 - 缺耐心 (业务评审会)
 - 缺时间 (乱尝试、过调参)



Alipay算法RD的工作模式

1. 平台化

- 数据地图（特征查询）
- ODPS + Sonic平台（特征清洗）
- PAI / Xlab 线下/上模型训练、数据分析平台
- UCT... 线上部署
- Alisis + 模型宝 监控报警

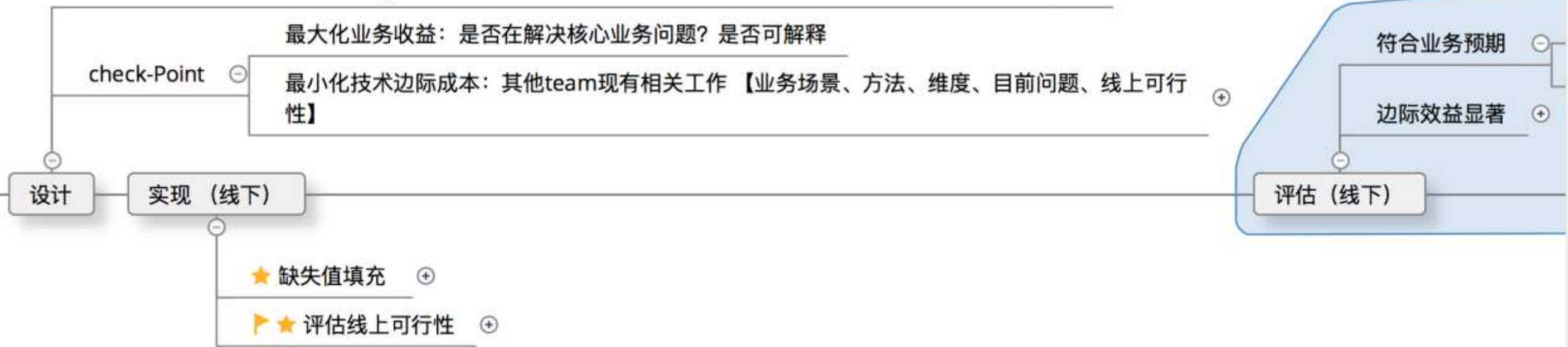
2. 技术溶于运营

3. 将特征做到极致



Journey & Alipay - 变量体系特征设计、开发、上线流程

启发 (e.g. 和运营讨论案例 & badcase & 变量体系)



低优先级

高优先级

技术选型 :

业务导向 : 在项目排期、业务预期等多方面约束下, 恰到好处的选型

优先级 : 项目最后一步

模型本身 : GBDT、LASSO

特征 :

价值 : 突破项目天花板的有效武器

业务导向 : 特征源自运营, 用特征充分刻画业务理解

方法论 : 变量体系、研发流程, 旧特征精细化、新特征拓展 (高维分类变量)

数据 : 数据是模型性能的上确界

训练数据 **一致对接** 线上预测的业务场景

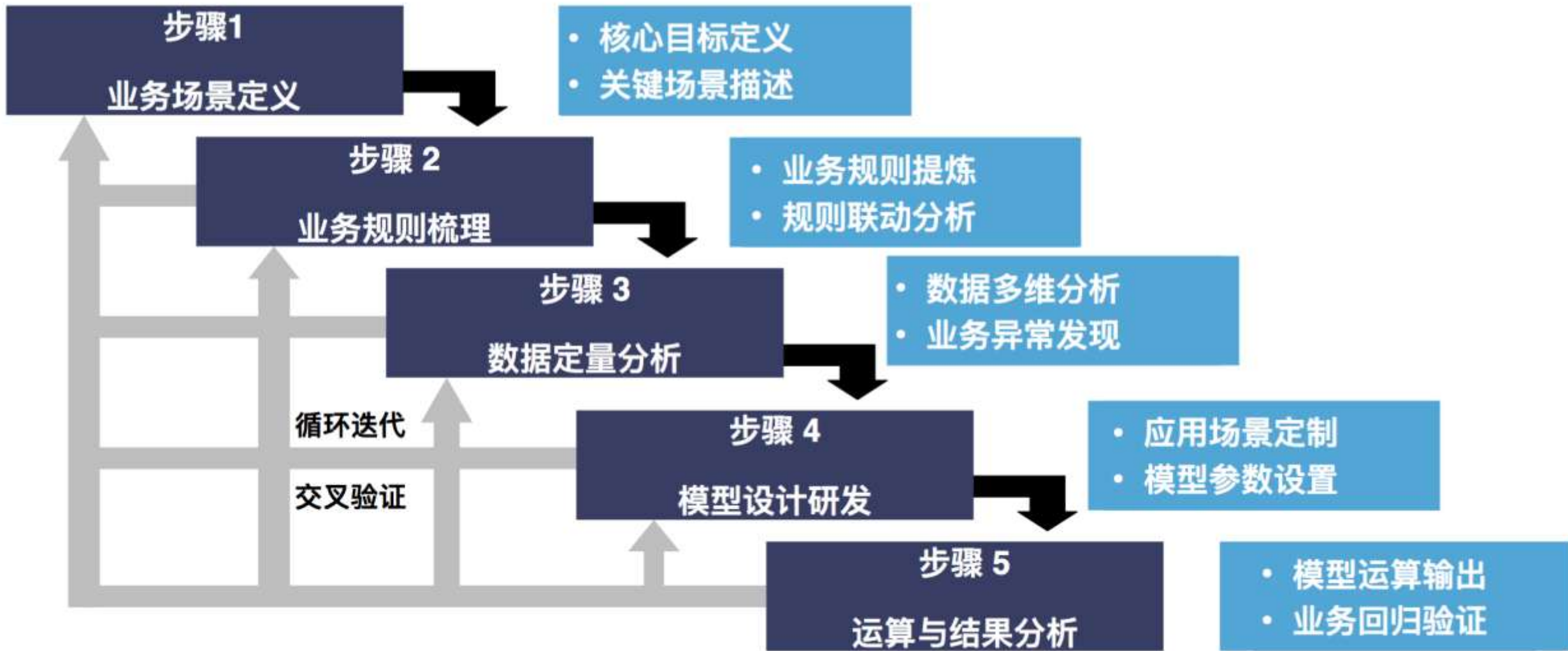
噪音的过滤 与 **99%准则**

业务 :

界定问题 : 业务核心目标、场景 (技术目标要倒推)

评估方案 ;

追求闭环 : 项目的输出是什么? 如何运营?



ROADMAP FOR ML PROJECT 📌

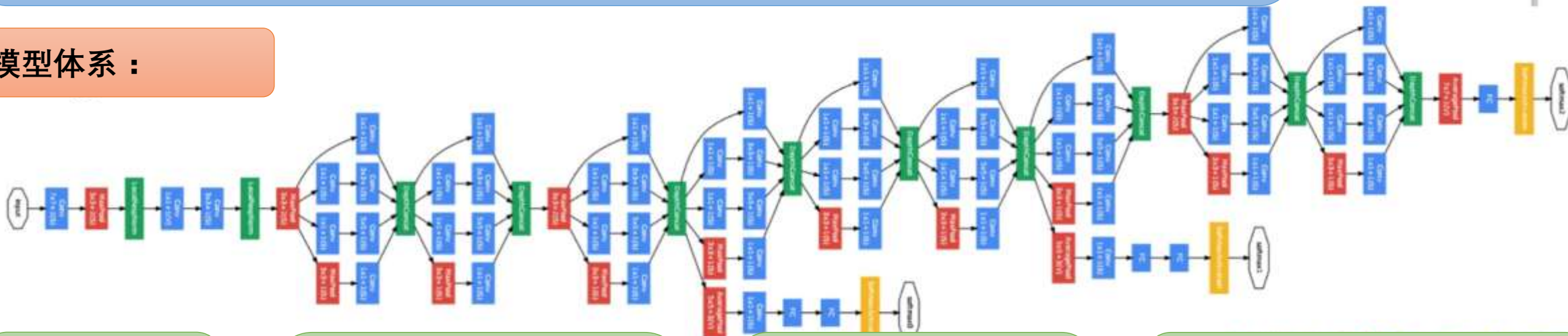
	业务 & 项目推进	特征 & 数据	模型
业务场景定义	访谈客户 ⊕ 资料收集		
业务规则梳理	内部讨论 ⊕	现有业务特征	在业务、数据、开发、排期的限制下选择恰当的模型方案
数据定量分析	数据分析, 验证技术方案有效性 访谈客户 ⊕	形成变量体系 特征线下开发、评估	如有必要, 用简单模型
模型研发、评估	访谈客户 ⊕ 协同开发 ⊕ 汇报	特征精细化 ⊕ 特征线上开发 ⊕	模型体系 复杂模型
线上开发、灰度	模型运营 ⊕	监控特征 ⊕	监控模型 ⊕



复杂业务：

- 同个业务方向，需要多个模型项目支撑 e.g. 风控体系、运力规划、尾部用户体验...
- 同个模型项目，服务多个业务方向 e.g. 外卖配送时间预估算法、配送调度算法、推荐排序

模型体系：



业务规则层：

1. 复杂业务高度抽象

基线层：

1. 刻画业务交集部分
2. 核心基础特征 e.g. UMID、关系、可信、基础行为
3. 模型深度、复杂、T+1

攻防层：

1. 刻画业务特性部分
2. 业务关注的长尾部分
3. Velocity

策略层：

1. 疑难杂症
2. 多目标联调

	架构	方法
策略层	风险识别: 70%+ HC & Effort; 1000+ Rules 风险控制: 人脸, 短信校验, 冻结...	案件复盘、播报 策略实验、数据分析
攻防层 KPI: 控制命中率, 优化召回率	手机丢失模型 → 批量登录模型 → 手机、pc木马模型 →	GBDT/RF
基线层 KPI: 控制召回率, 优化命中率	<p>账户 (主体/客体) → 介质 (手机/sim, 银行卡, 设备) → 地点 (城市, wifi, 基站) → 个体行为体系 (1. 金额, 2. 时段, 3. 交易物品, 4. ...)</p>	RFM Logistic Re WOE 无监督学习
技术挑战	每天 1.5次交易都需要实时、高召回率、高命中率的监控	



	职能	模型	项目
策略层 业务方向协同	面向效率、体验、极端case等非常规机器学习问题，做贴近业务的策略模型	时间与用户体验策略 效率、体验最优预计送达时间 策略补时 <ul style="list-style-type: none"> 金额 出餐慢 负载 午夜 疑难杂症 	预调度、时间与用户体验 极端case优化 协同优化
攻防层 极端case补时	面向“突变”pattern，实时优化 (online-learning)	送达时间实时优化 未来N分钟时间、单量	实时优化 未来Nmin指标预估
基线层 历史pattern&复杂模型	面向对历史pattern的深度挖掘学习	送达时间 入离店时间差 交付时间 出餐时间	三段论模型 骑手轨迹 入离店时间差 交付时间预估



携程技术中心



IT大咖说
知识分享平台

THANK YOU!

Q&A