

深度学习在外卖的应用

刘怀军@美团外卖 201706



1.

- 业务介绍

2.

- DNN在图像技术应用

3.

- DNN在预估模型应用

4.

- 总结

O2O行业

团购

美团网
meituan.com

大众点评
dianping.com

百度糯米

窝窝团
55tuan.com

拉手网

点评

大众点评
dianping.com

OpenRice
開飯喇

外卖

淘点点

我有
外卖
Yes,
I Deliver

锦食送
JinShiSong.com

饿了么

生活半径
SHENGJING.COM

外卖单

美团外卖

waimaidan.com

百度外卖

Dianwoba
点我哪巴

零号线
ZINEO.COM

外卖超人

易淘食
就食易淘，易淘时吃

预订

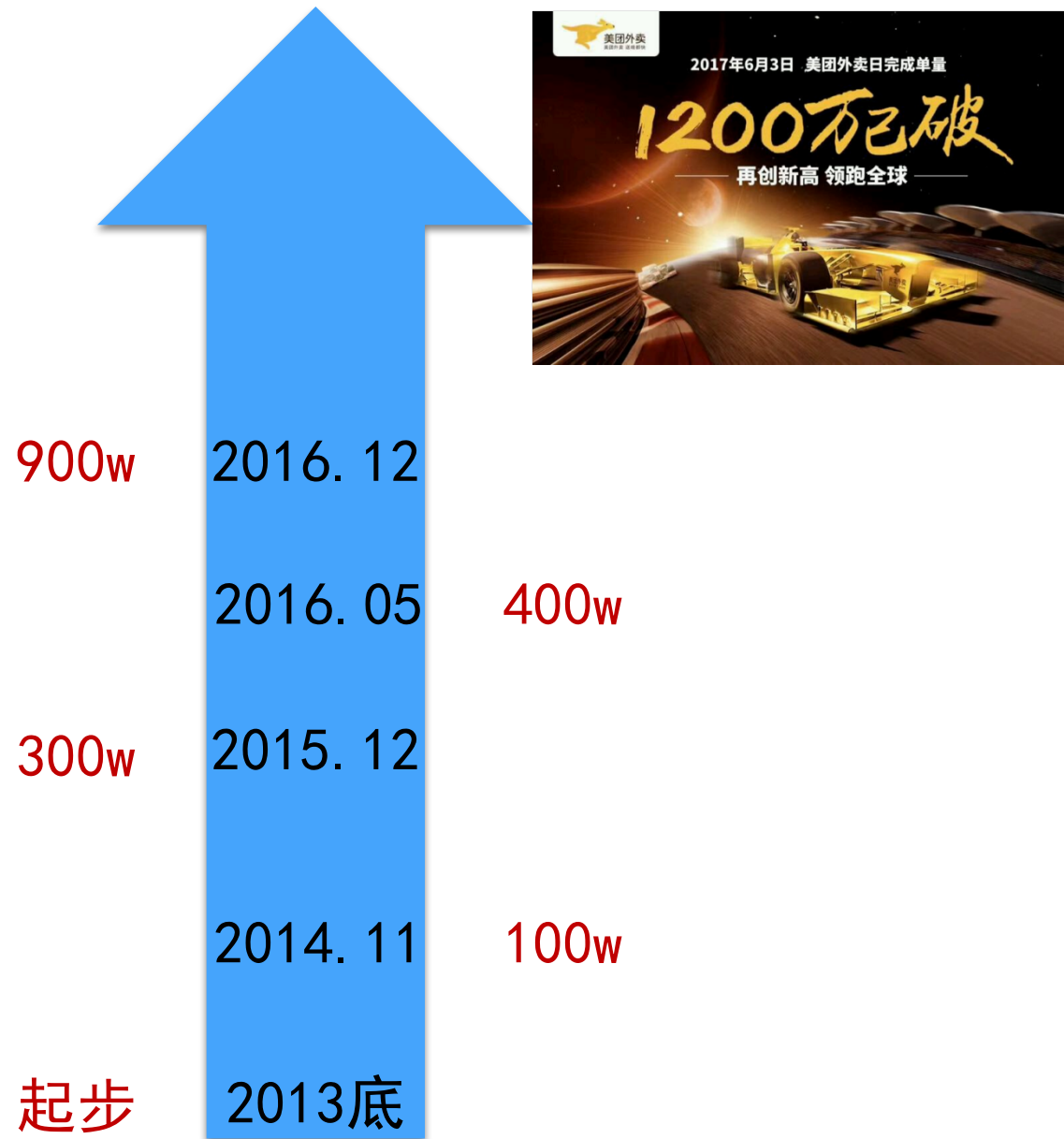
EZTABLE

订餐小秘书
www.xiaomishu.com

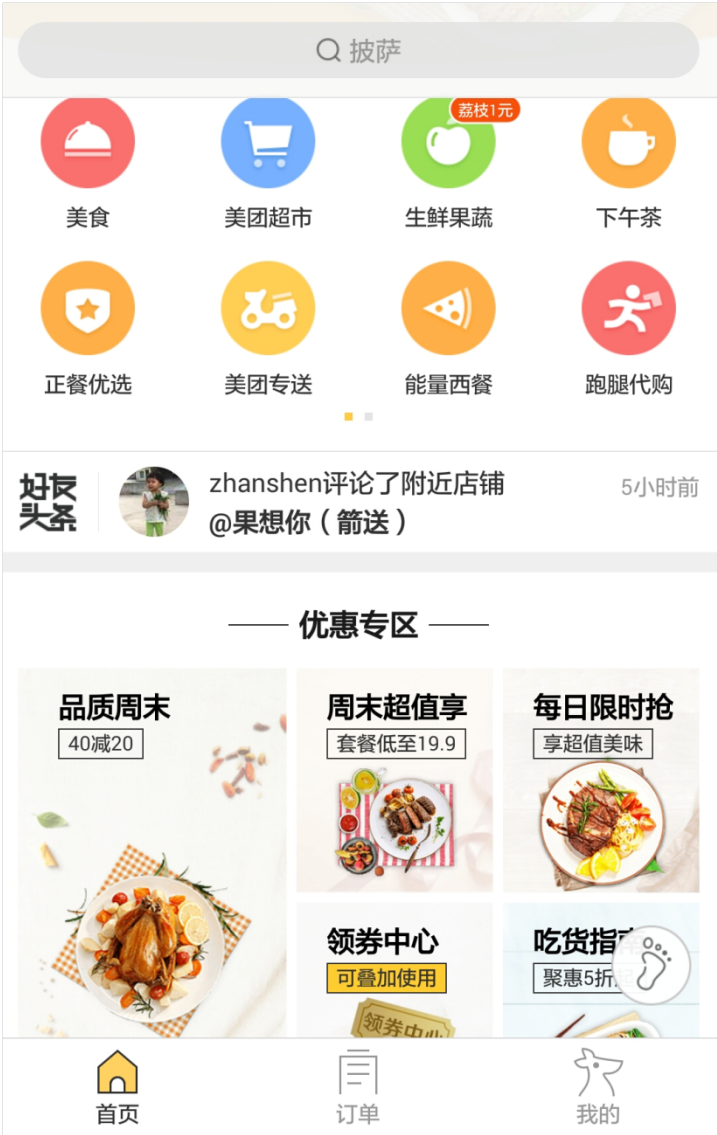
排队网
www.paidui.cn

宴请网
www.yq940.com

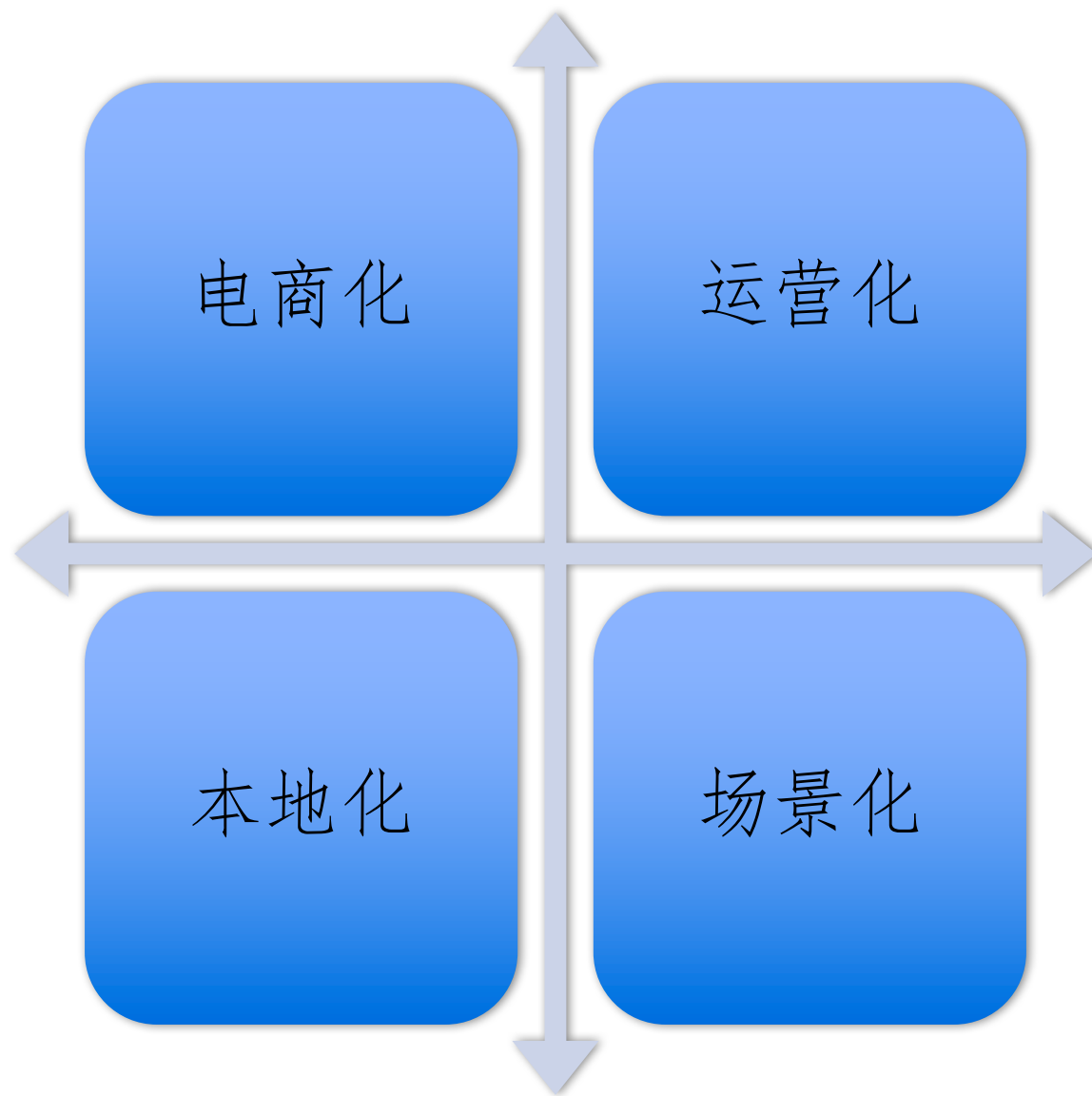
美团外卖



产品形态

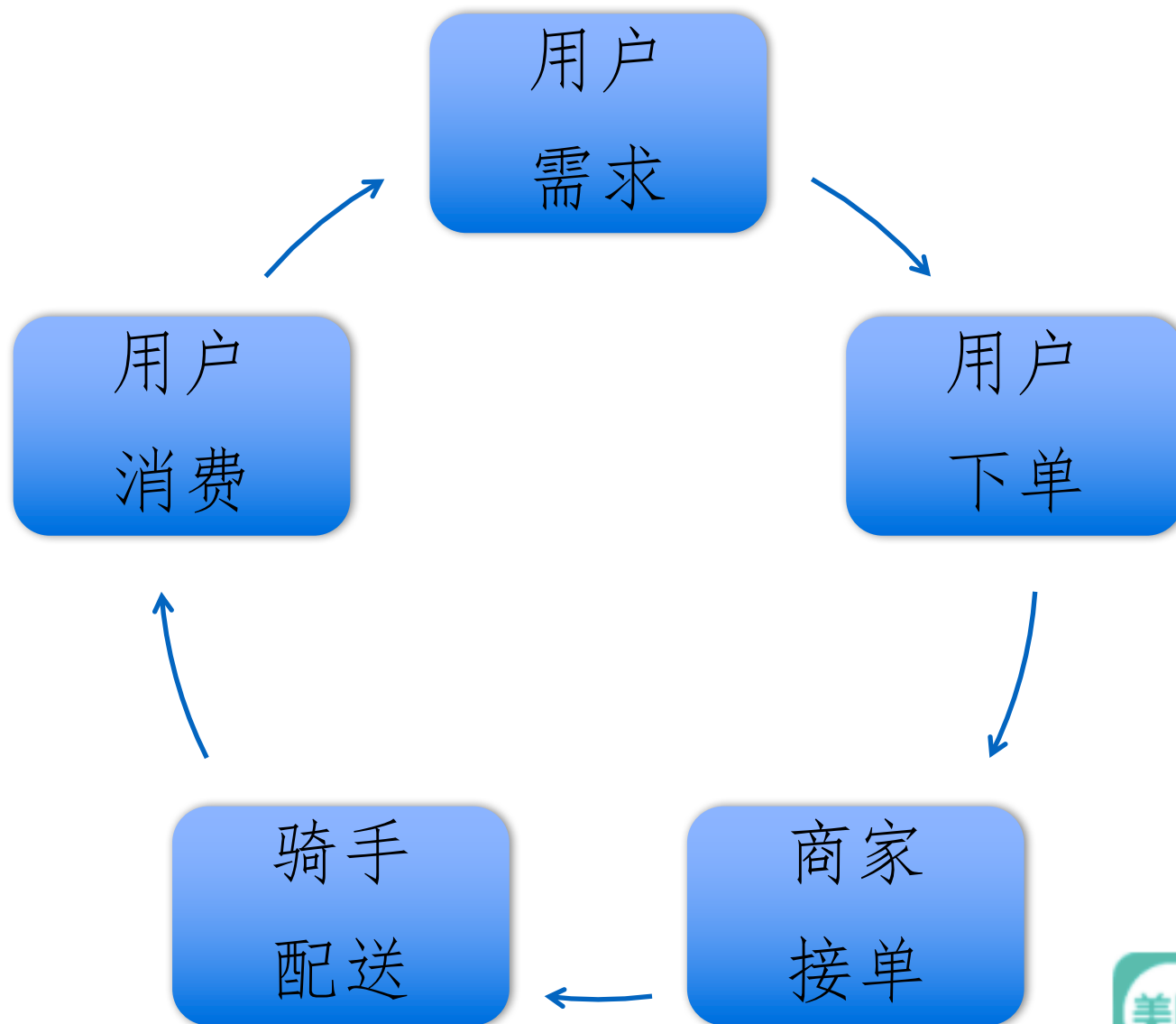


业务特点

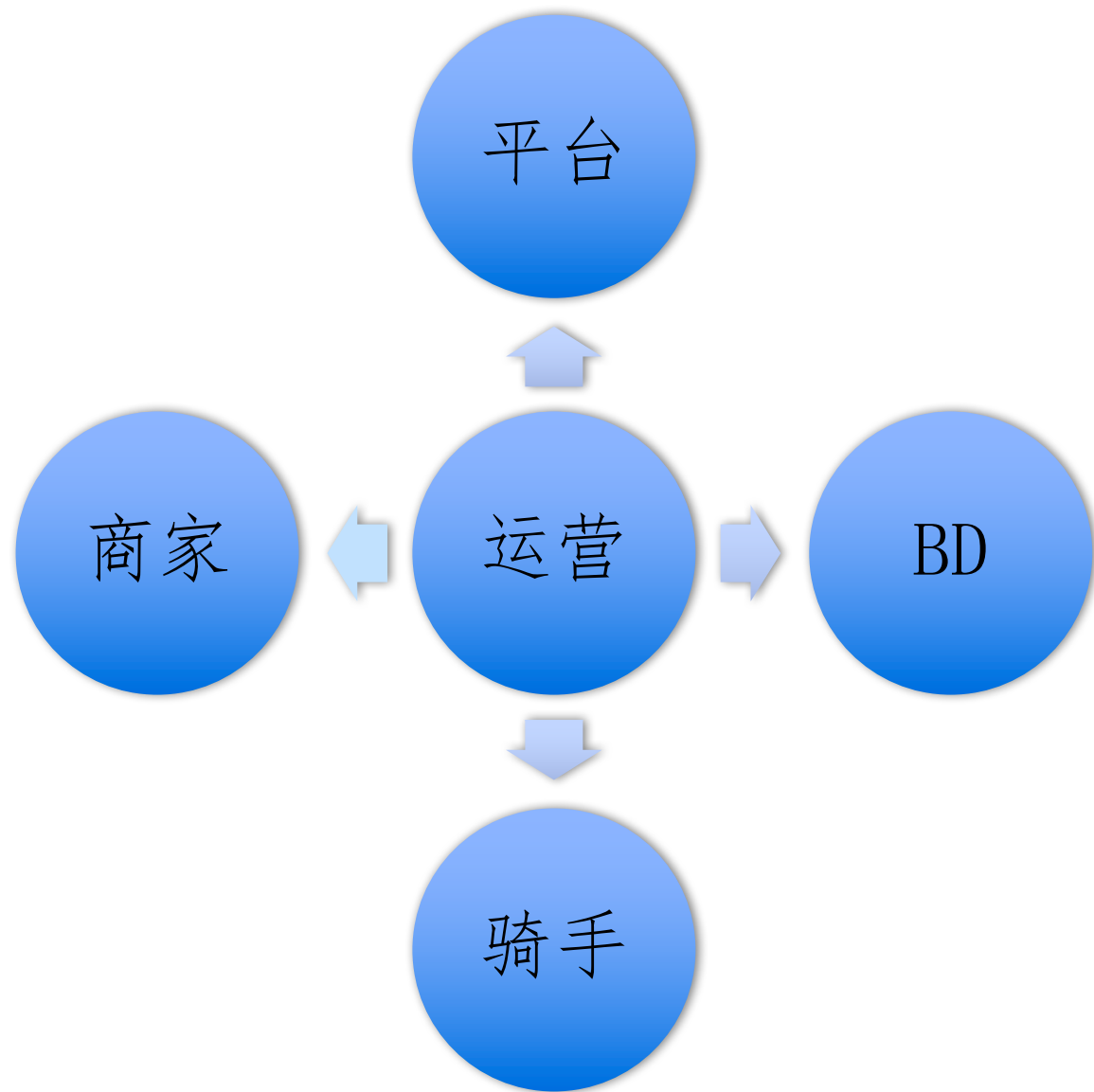


业务特点-电商化

- 线上线下结合
- 交易闭环化

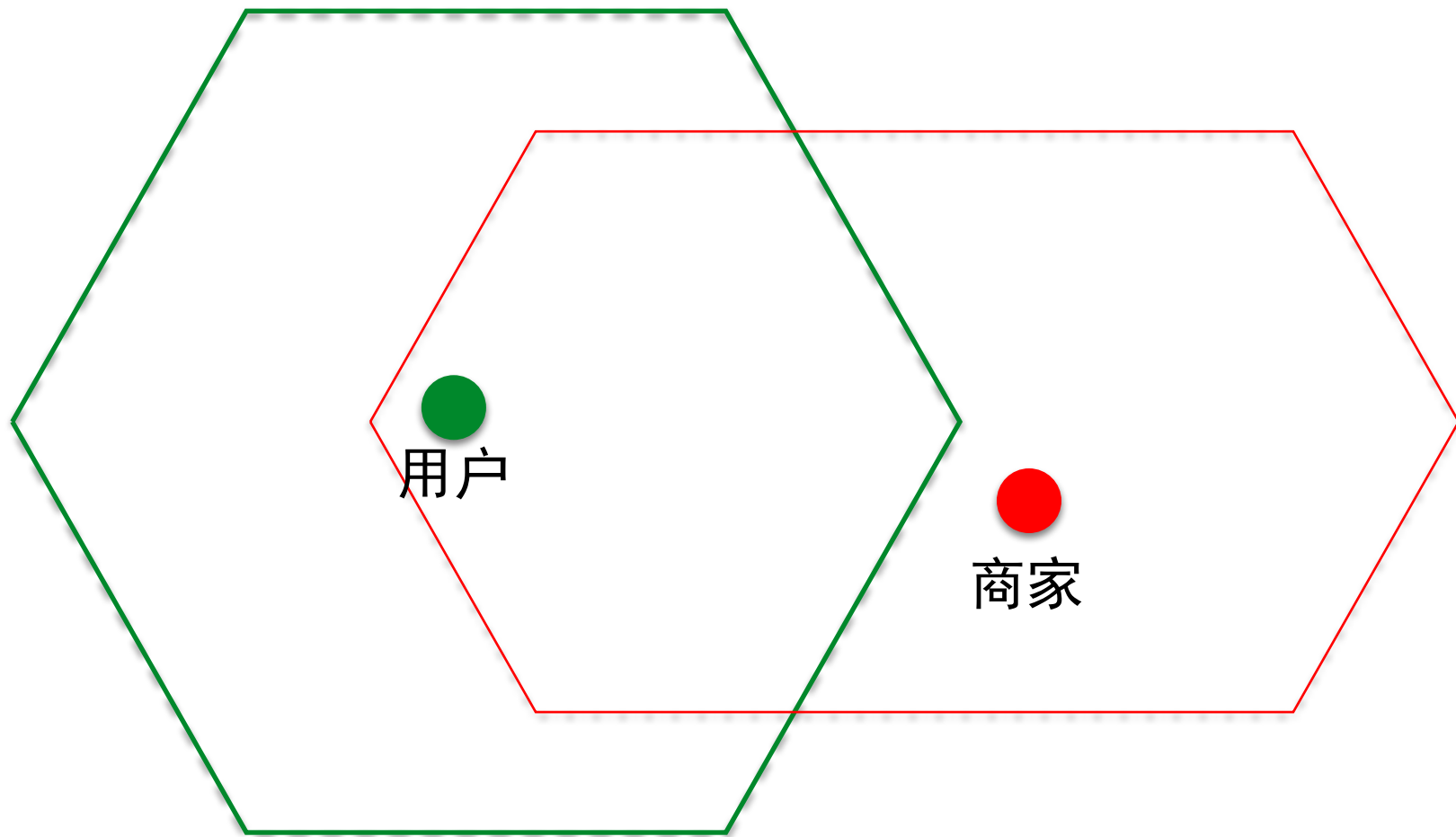


业务特点-运营化



- 供给内容形式多元
- 运营能力参差不齐

业务特点-本地化



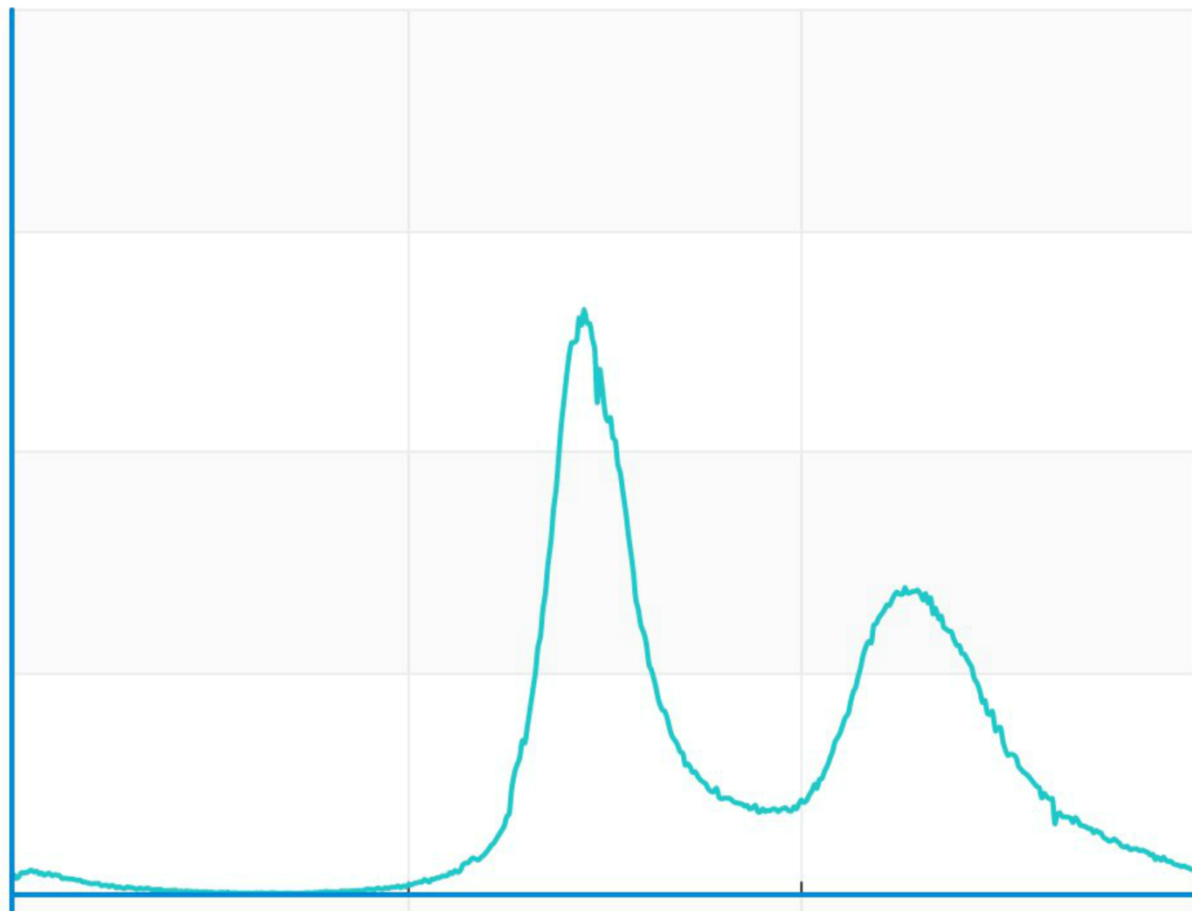
- 供需范围被约束
- 供需关系动态化

备注：图中多边形仅用于举例

业务特点-场景化

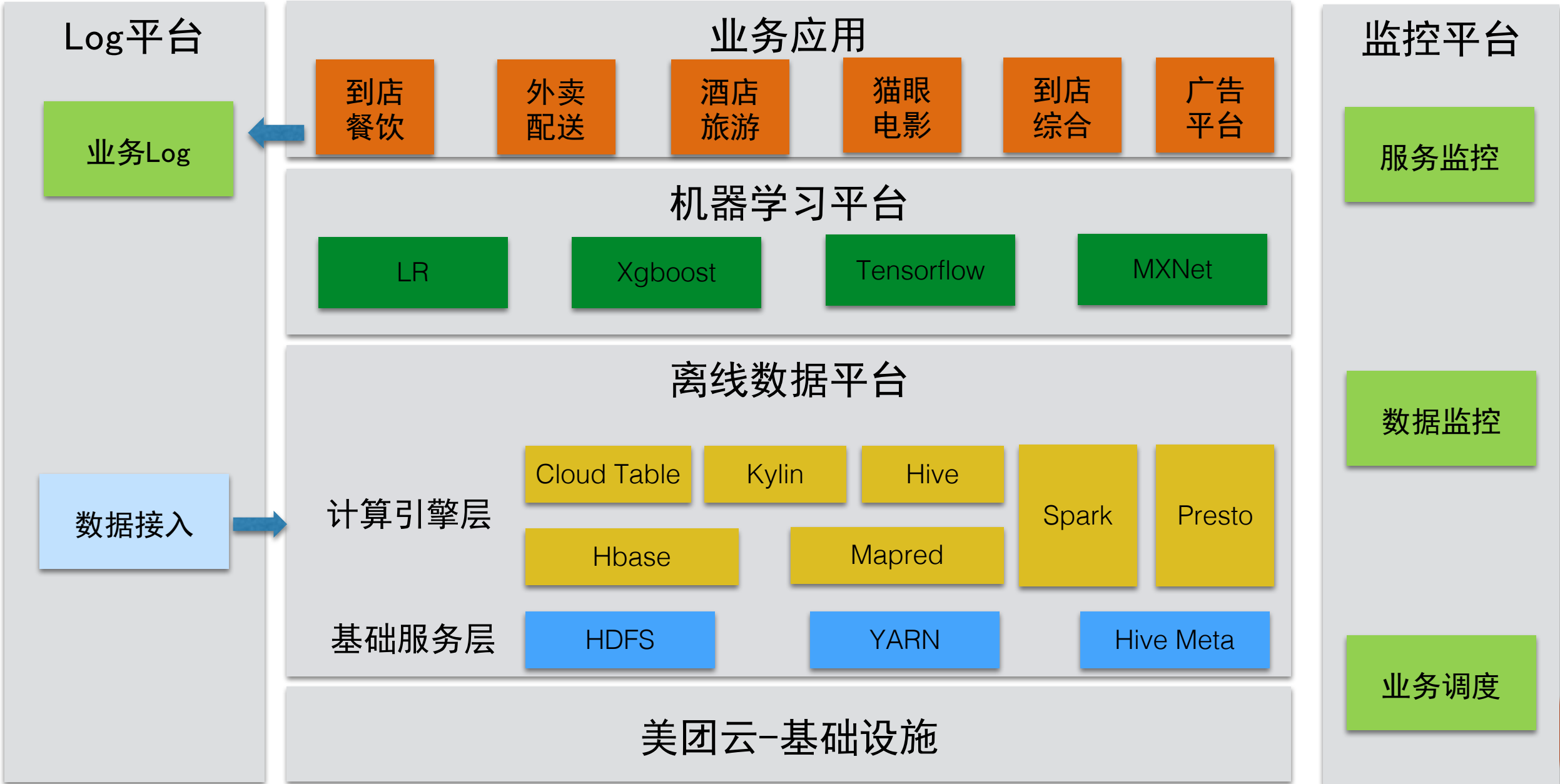
- 时段化差异显著
- 地域性差异显著

美团外卖单天订单趋势图



- 个性化程度高
- 精细化需求大
- 多元化特点强
- 实时化强度高

美团点评大数据平台



大纲

1. • 业务介绍
2. • DNN在图像技术应用
3. • DNN在预估模型应用
4. • 总结

案例1：图像质量技术

背景



VS



- 用户评判标准主观
- 用户关注点差异大
- 人工标注成本高

多维度特征设计

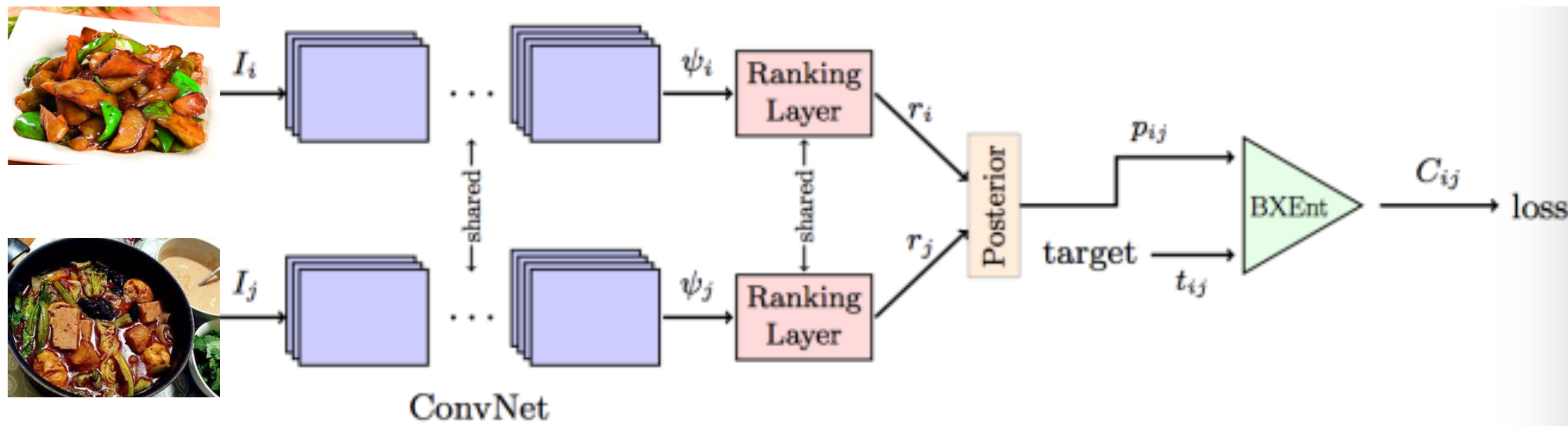
- 美学特征
- 基础特征

自动化数据标注

- 图片来源
- 用户行为

深度学习建模

方案



高阶特征

- 美感
- 吸引度
- 品类
-

基础特征

- 色彩
- 锐度
- 对比度
-

pairwise

- 不直接标注样本tag
- 进行两个图是质量对比
- vgg+ranknet方法

- 用户体验提升
- 结果点击率提升

案例2：OCR技术

自动审核

- 营业执照
- 餐饮许可证
- 商家其他证件

自动录入

- 饭店菜单识别
- 饭店招牌识别

- 图片成像复杂
- 文字构成复杂
- 图像背景复杂

方案-文字检测

FCN

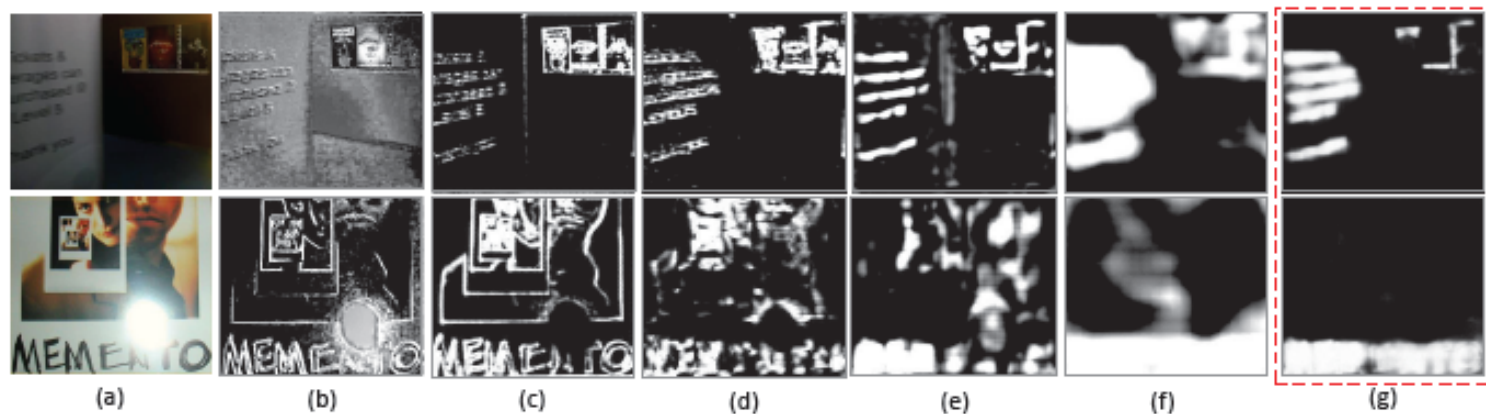
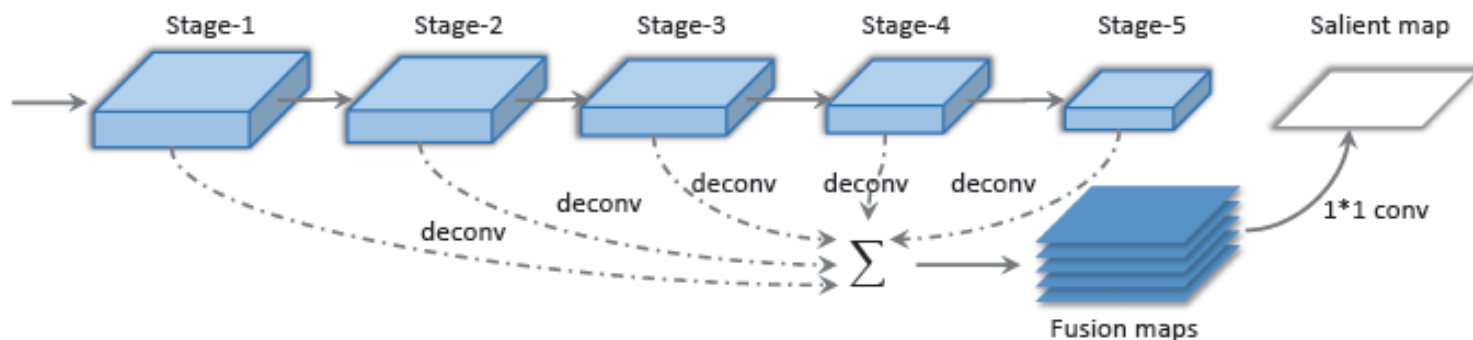
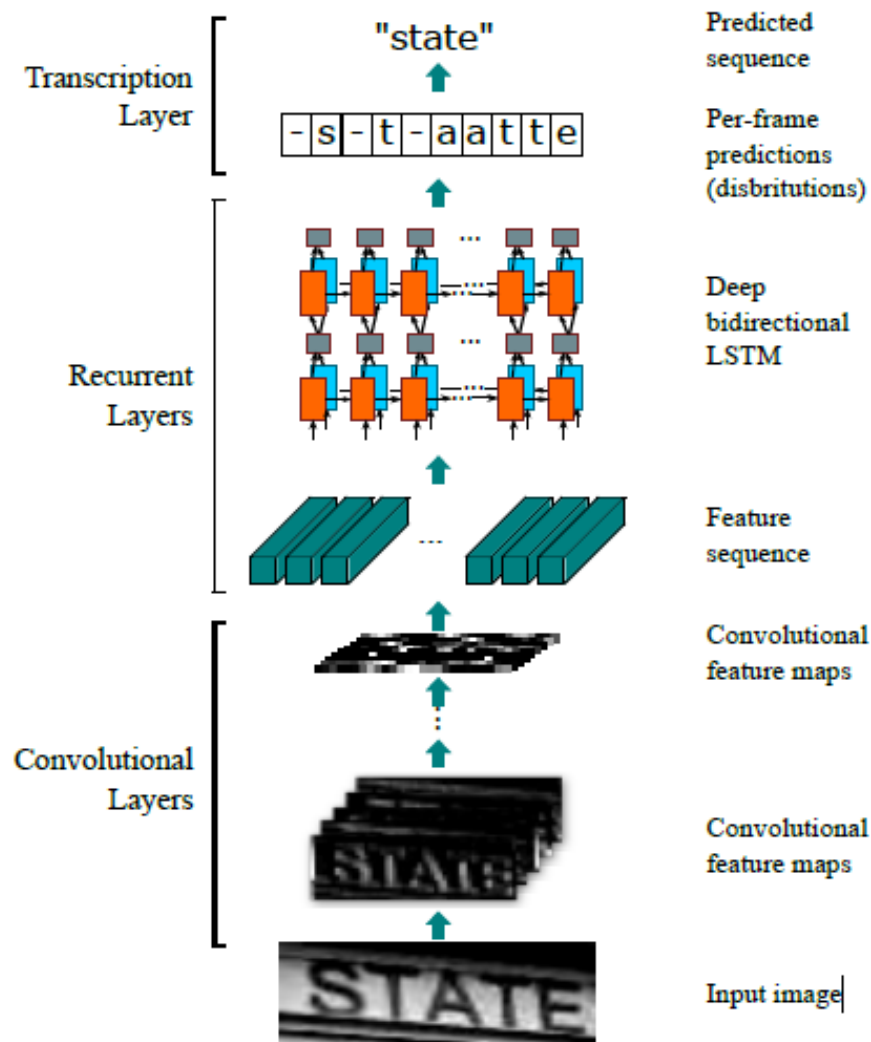
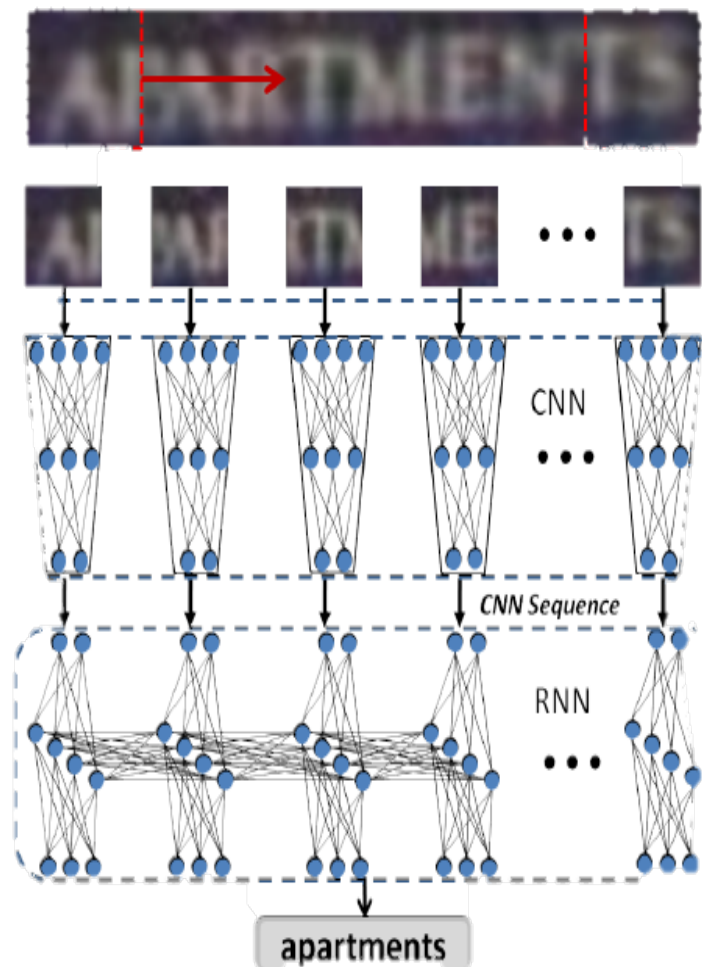


Figure 2. The illustration of feature maps generated by the Text-Block FCN. (a) Input images; (b)~(f) The feature maps from stage1~stage5. Lower level stages capture more local structures, and higher level stages capture more global information; (g) The final salient maps.



方案-文字检测识别

CNN特征+BLSTM+CTC 序列识别模型

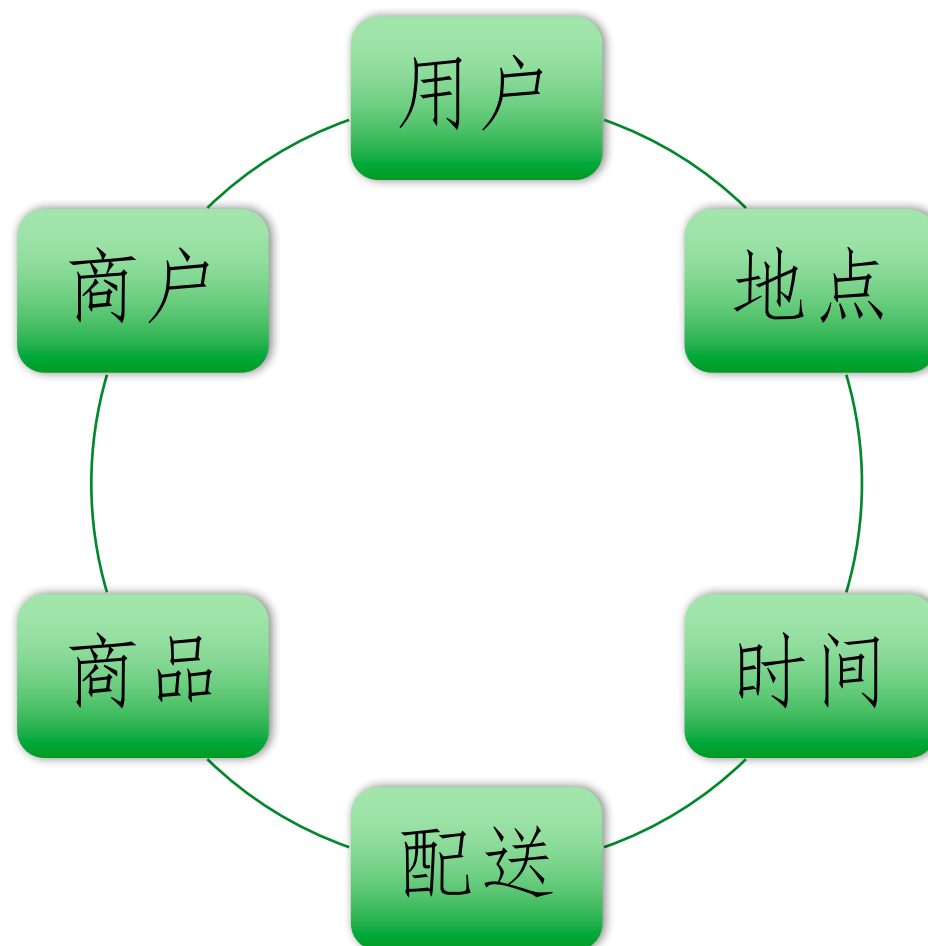


- 人工化→自动化
- 运营效率提升

大纲

1. • 业务介绍
2. • DNN在图像技术应用
3. • DNN在预估模型应用
4. • 总结

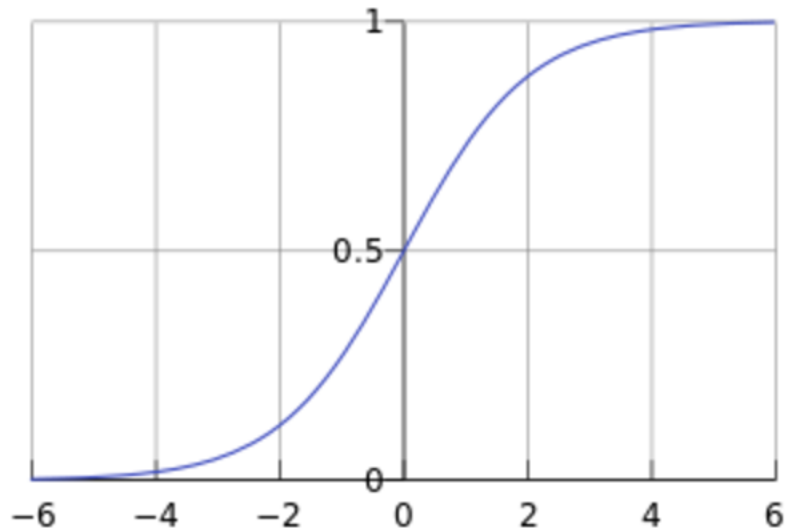
背景-业务模型



背景-特征工程挑战

- 特征构造成本高
- 特征选择开销大
- 依赖人工经验重

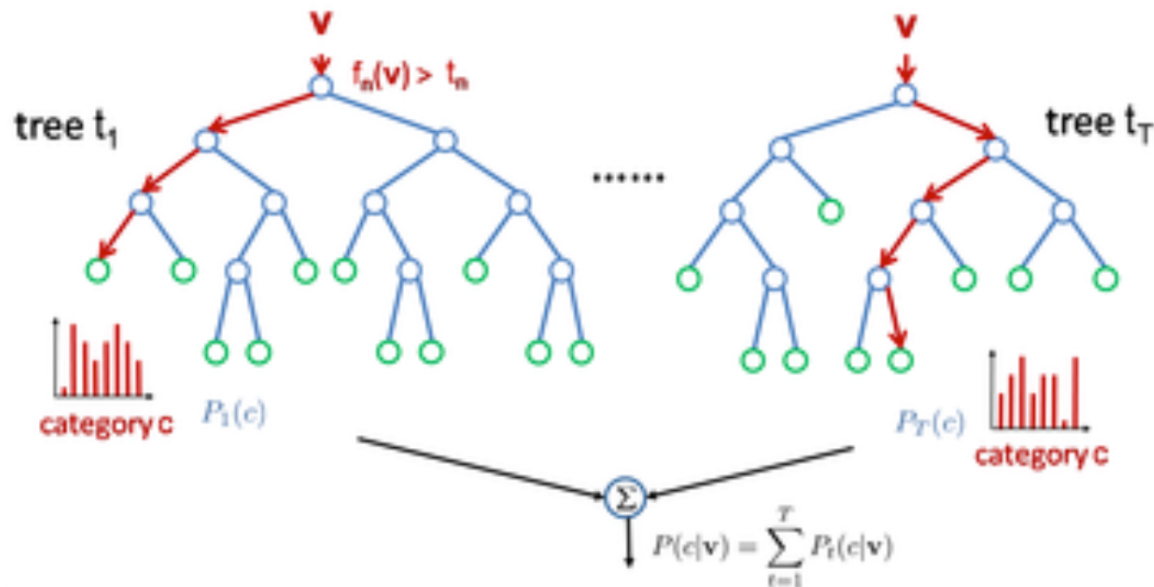
背景-传统模型问题



$$h_{\theta}(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$

LR模型

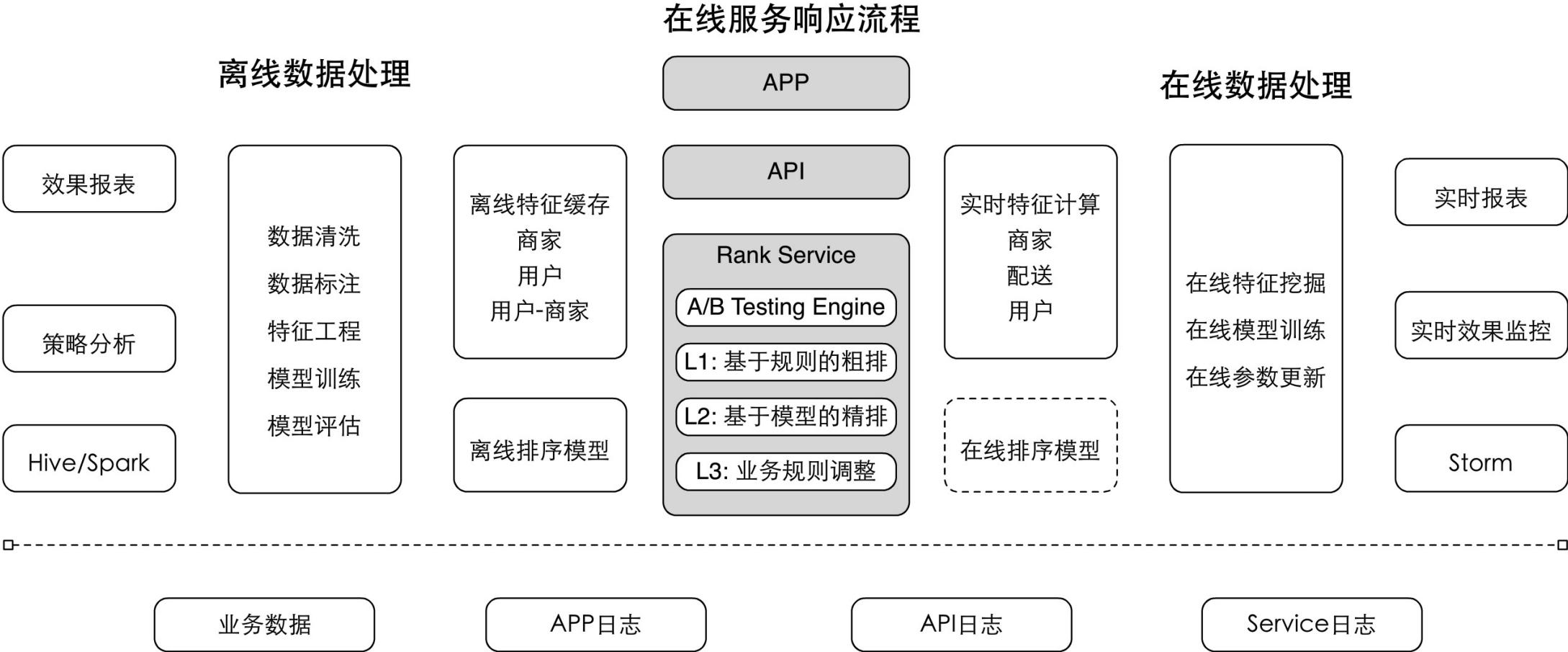
- 非线性拟合能力受限
- 需要手动组合特征，成本高



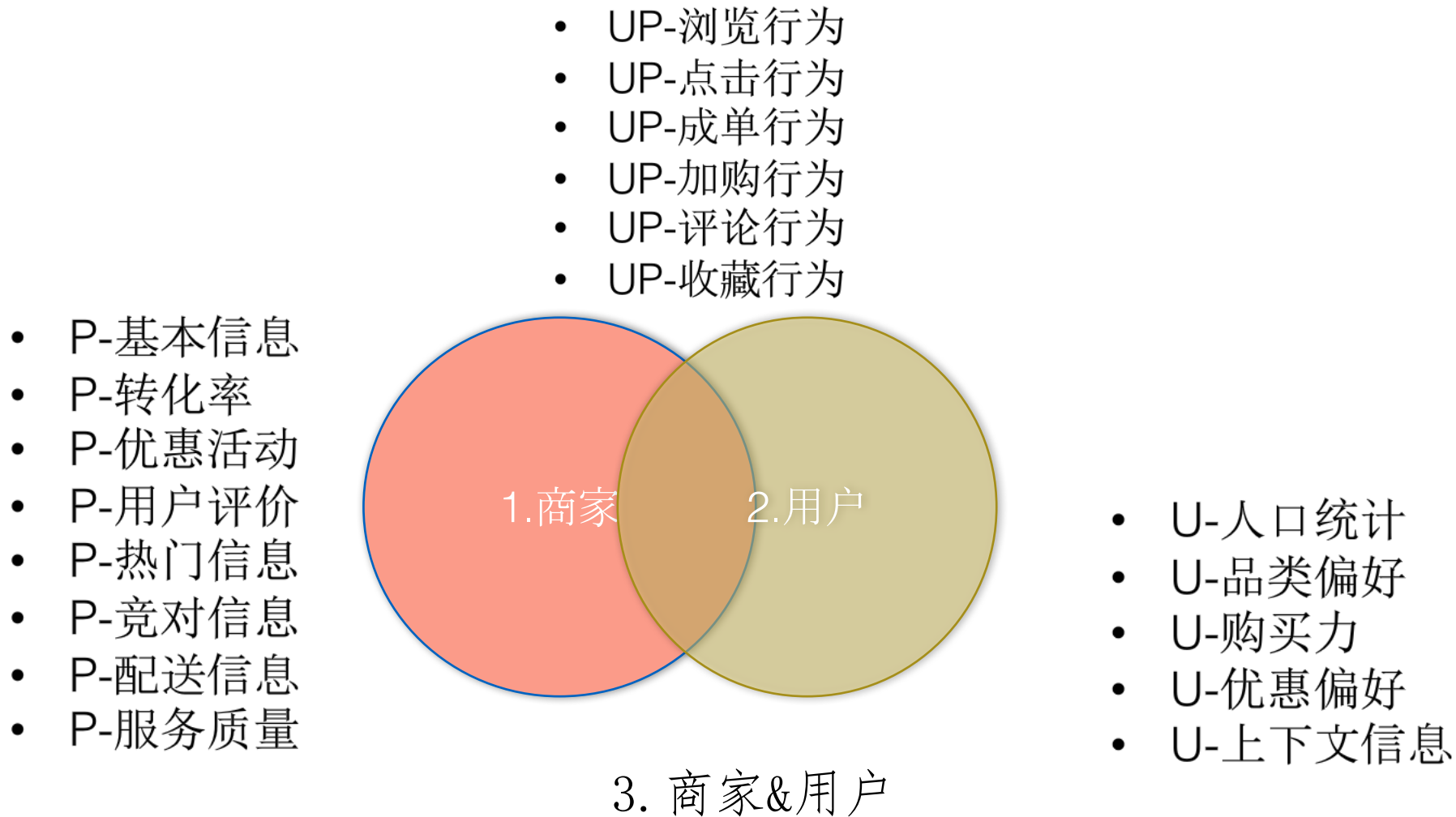
树模型

- 不适合高维、稀疏特征
- 无法对文本语义直接建模

方案-系统架构



方案-特征体系



方案- Wide&Deep Model

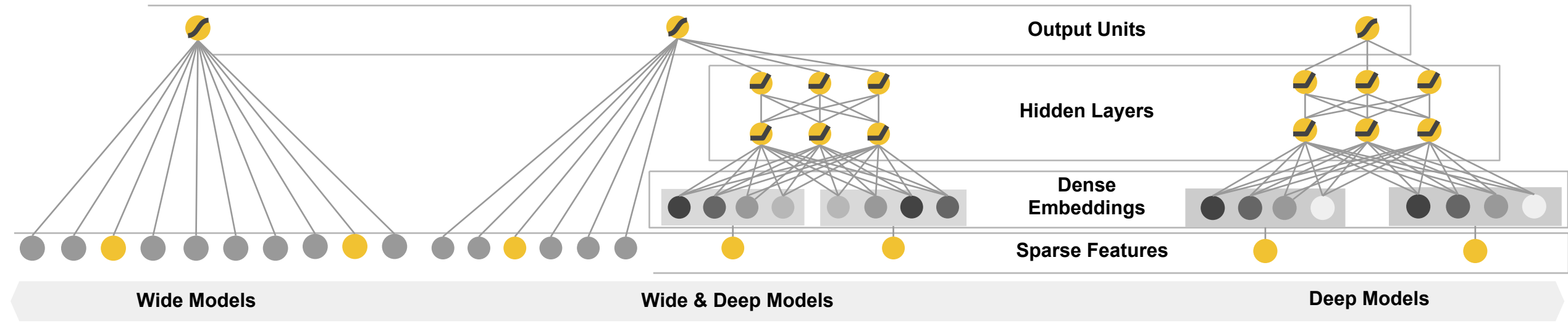


Figure 1: The spectrum of Wide & Deep models.

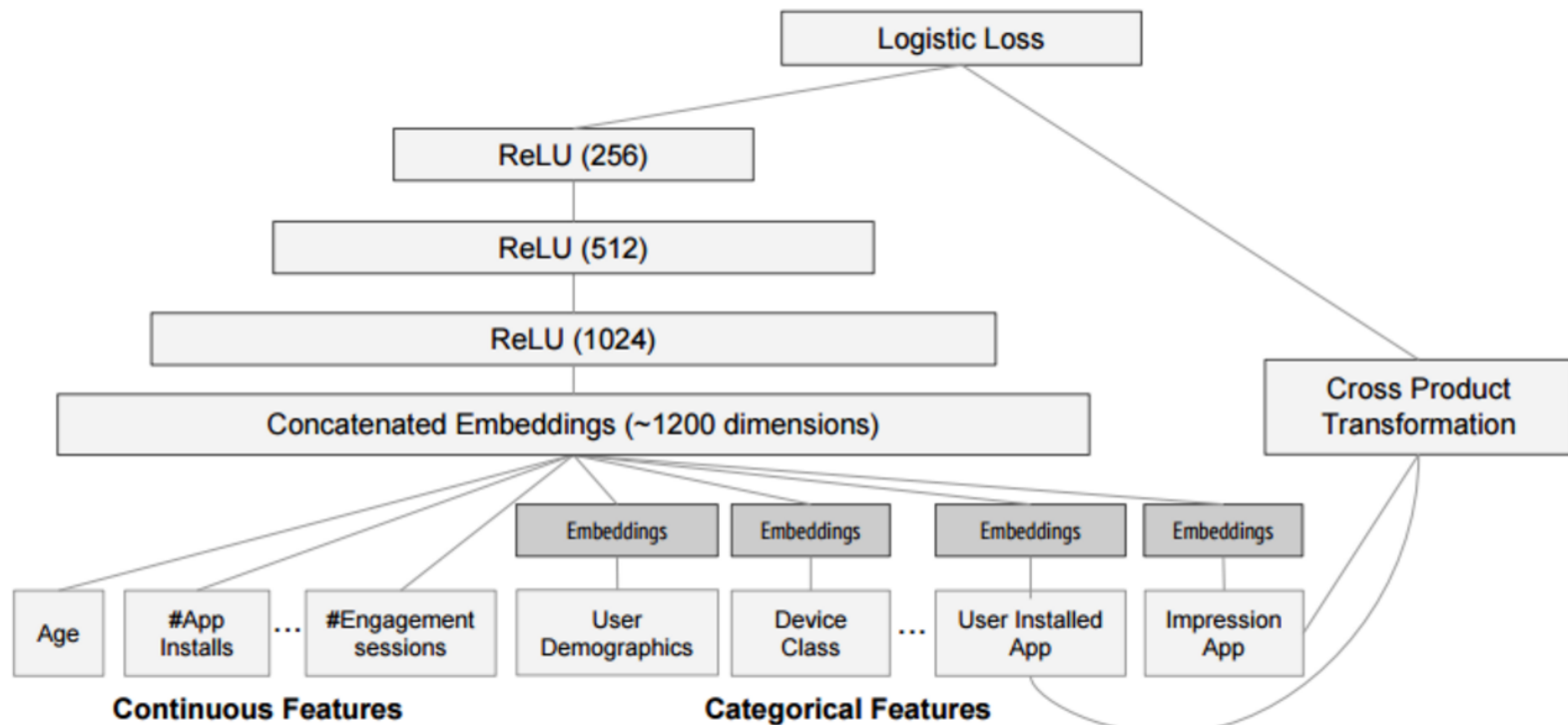
方案- Wide&Deep Model

Wide Model

- LR-线性model
- 包含组合特征
- 特征one-hot到binary特征

Deep Model

- NN网络
- 输入高维、稀疏feature, eg: category
- 特征泛化: embedding到低维实数向量



方案- Wide&Deep Model-离线流程

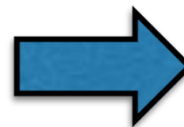
样本处理

- 样本清洗
- 爬虫过滤
- 无效加载

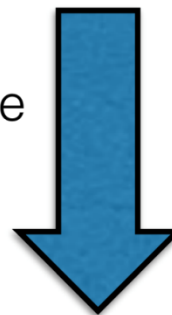
采样优化

- 正负样本的选择
- 高频用户/商品

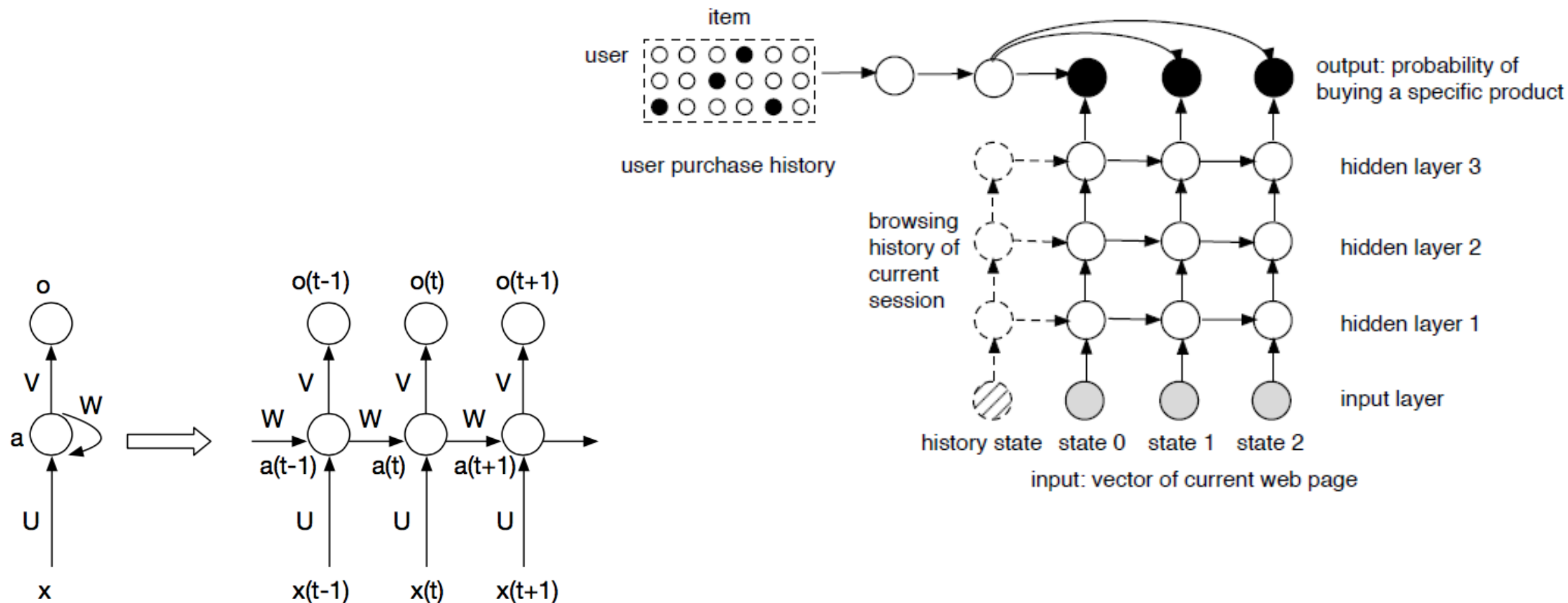
最后点击/下单



skip above
+2

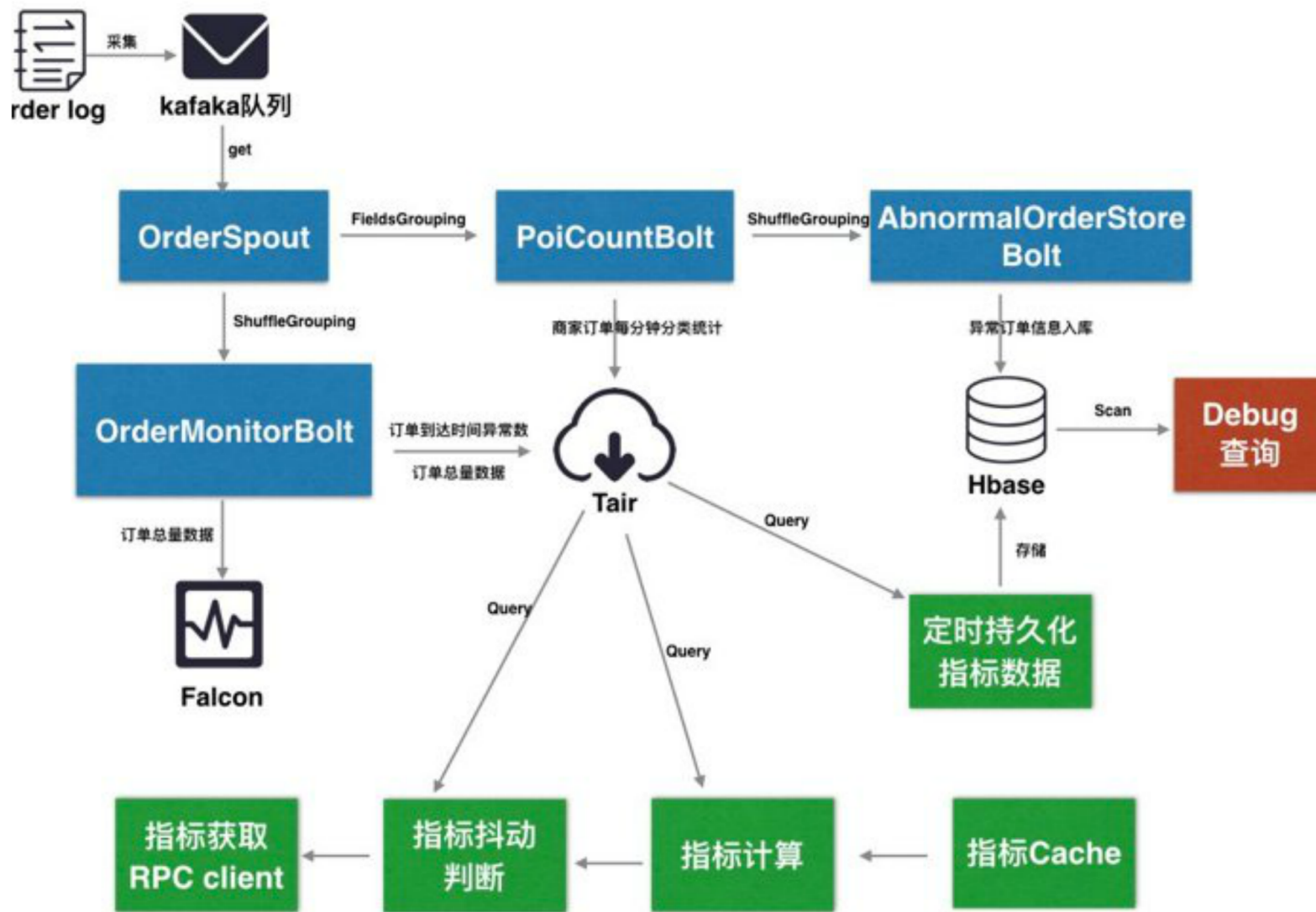


方案- RNN Model



Recurrent Neural Network

方案-实时化架构



方案-RNN Model

输入

- 用户 session 的浏览序列 $in = \{i_1, i_2, \dots, i_m\}$
- 元素 i_x 对应输入层一个状态, one-hot encoding

输出

- 用户可能结果概率 $out = \{$
- 概率最大元素为预测结果 $\left. \begin{matrix} o_1, o_2, \dots, o_m \end{matrix} \right\}$

超参

- 序列长度
- 按一定时间窗约束



- 推荐精准度提升
- 用户体验提升
- 平台转化效果提升

大纲

1.

- 业务介绍

2.

- DNN在图像技术应用

3.

- DNN在预估模型应用

4.

- 总结

总结

- 技术对外卖业务至关重要
- 深入理解业务+技术精细化调优
- 不断尝试+总结分析

等你来

算法数据大牛
共创最强外卖大脑!

人工智能

外卖领域人工智能起航，创造未来

岗位：图像/语音/NLP算法

精准营销

让营销和广告投放更智能

岗位：广告算法/架构

搜索推荐挖掘

下半场精耕细作，需要千人千面

岗位：搜索/推荐/挖掘算法和架构

大数据

外卖掘金风口，大数据助力

岗位：数据开发



简历投递：wangxiaoxi02@meituan.com

业务介绍:

外卖作为到家场景核心业务，2017年中日订单突破1200万，跃居市场第一。

互联网下半场开始，数据算法驱动业务精耕细作，实现高速增长。人工智能起航，语音、图像及语义黑科技，实现外卖颠覆性突破。数亿用户和千万商户高频度、类型丰富的行为数据，等你来挖掘淘金。

加入我们，一起创造外卖领域数据、算法和人工智能的辉煌！

美团外卖数据团队招聘

Thanks

Q&A