

深度学习在推荐中的实践

陈文石@美团点评



陈文石

推荐平台组，美团点评/点评平台搜索推荐智能技术中心

百度自然语言处理部

大众点评算法团队 商户推荐/关键词搜索/B端数据挖掘

美团点评搜索推荐智能技术中心/推荐平台组

- 推荐引擎概述
- 点评推荐平台
 - 业务背景
 - 平台架构与演进
 - 策略迭代
- 深度学习及在推荐中实践
 - 项目介绍
 - 深度学习推荐
 - 深度学习应用
- 思考及展望

推荐引擎概述：为什么需要推荐？



- 大数据时代的悲伤
- 海量的碎片化信息，大脑处理信息能力有限
- 随之而来的选择恐惧症
- 推荐的使命
- 根据用户偏好，发现用户需求，协助用户决策
- 形成高效的信息分发方式

提升订单转化



35%的GMV来自于推荐

提升信息分享传播效率



DAU 13亿，5200亿美元

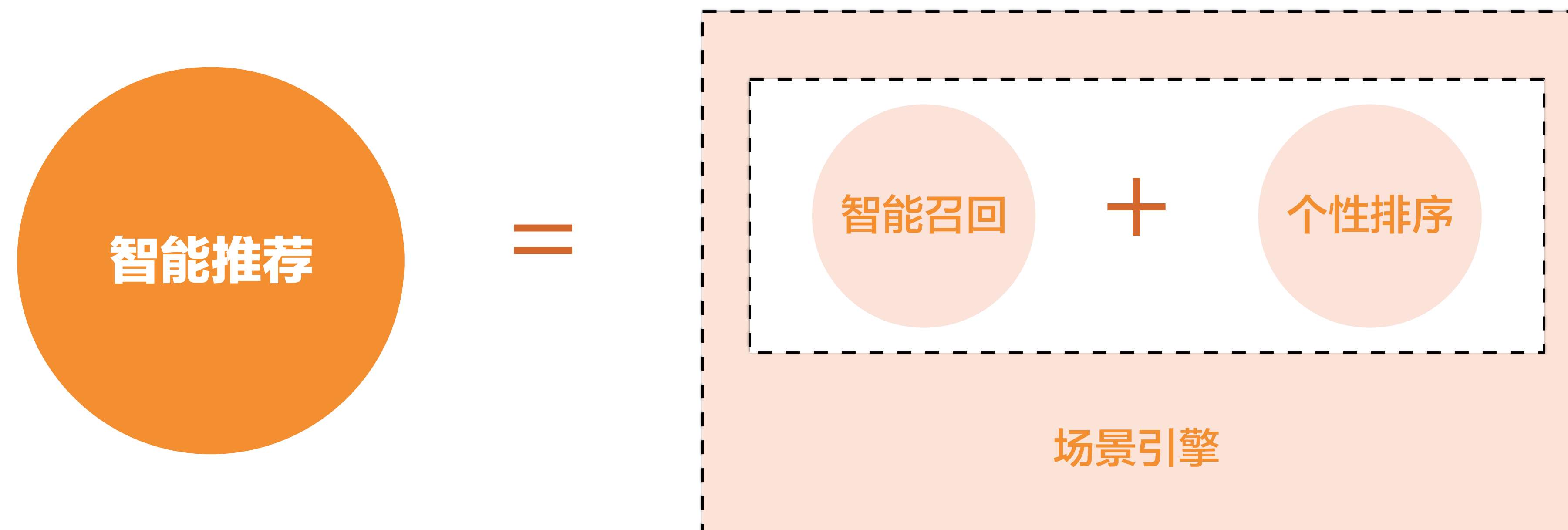
提升内容分发效率

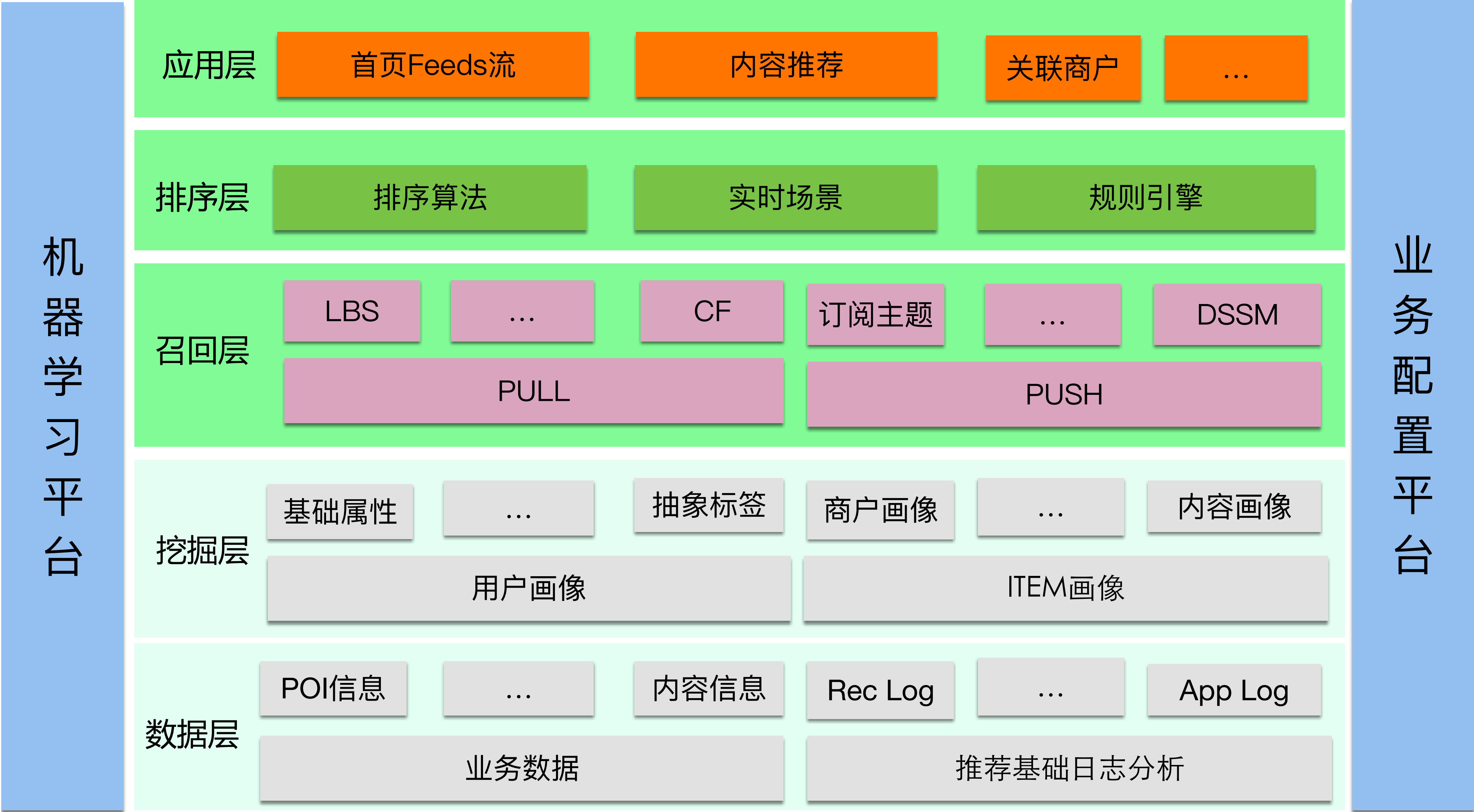


DAU 1亿，用户时长74分
钟

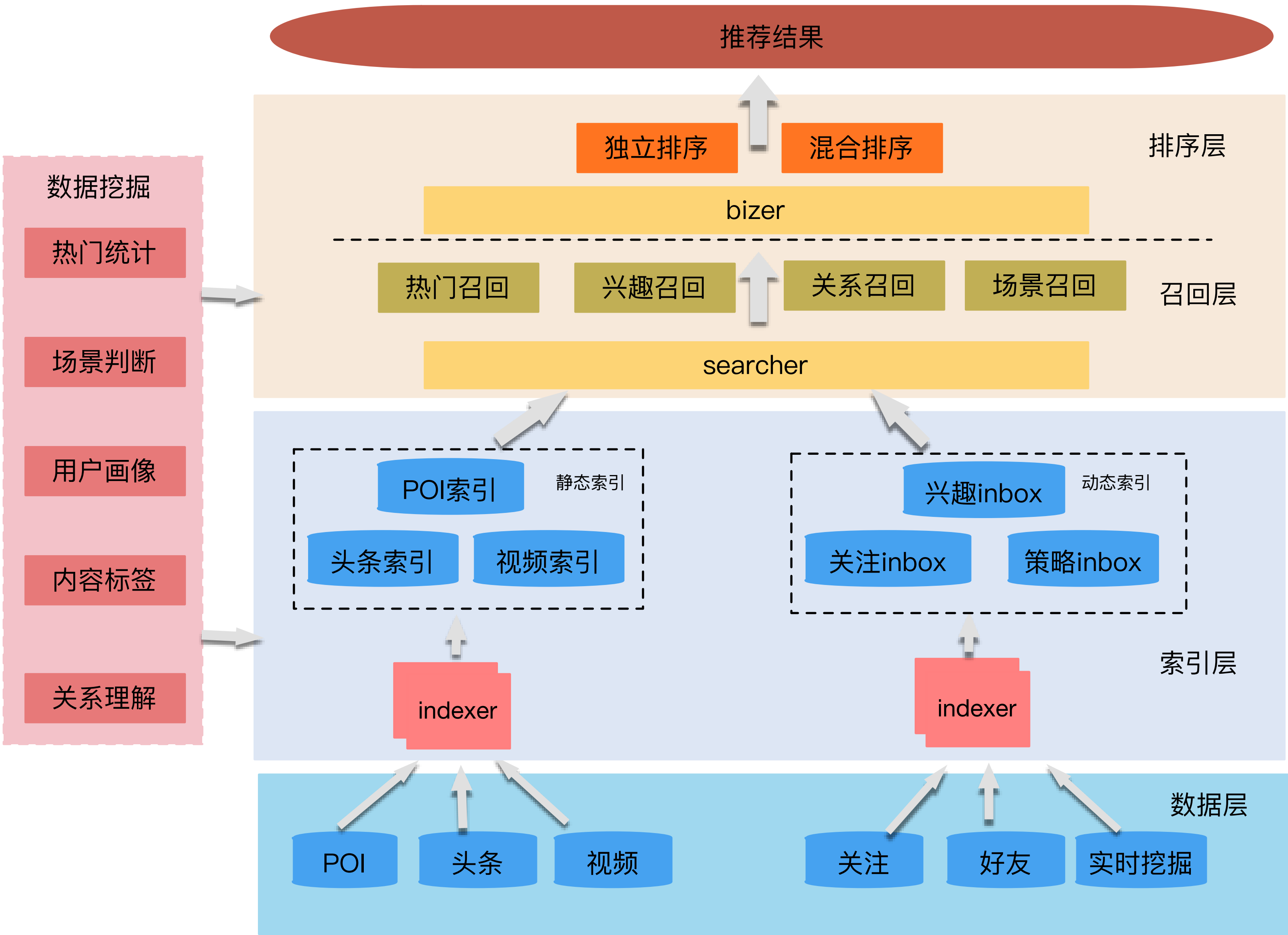
点评推荐平台：业务背景







点评推荐平台：服务架构

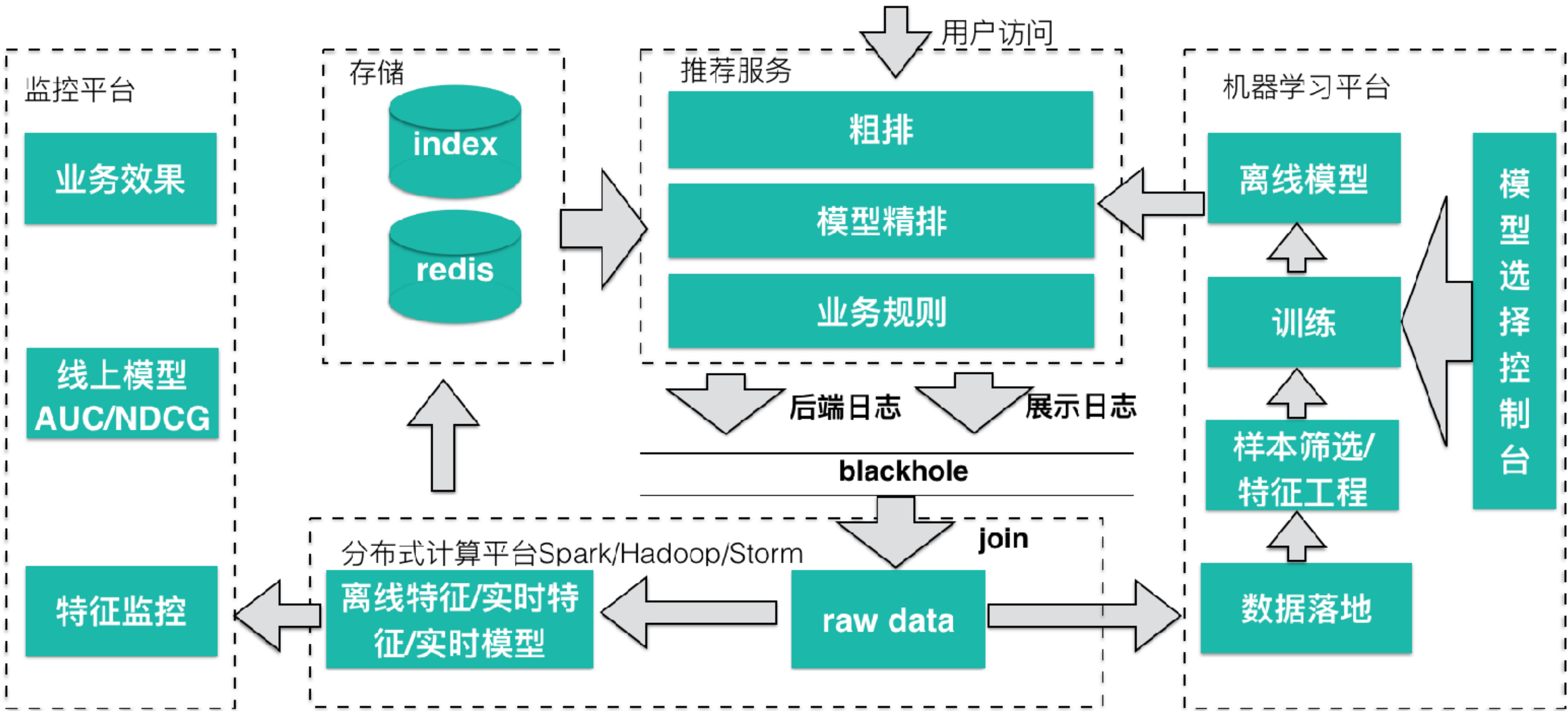




架构演进：从Pull到Push（以空间换时间）

- Pull方式适合LBS、CF等静态召回策略
- Push方式适合用户订阅、好友阅读、近线算法等召回策略

点评推荐平台：排序架构



召回策略



排序模型



- 背景

- 2017年7月，首页进行大改版，从运营与产品驱动转变为大数据和机器学习算法驱动，将运营位置融入Feeds流

- 优化目标

- 提升信息分发与内容消费的效率，优化用户体验优化

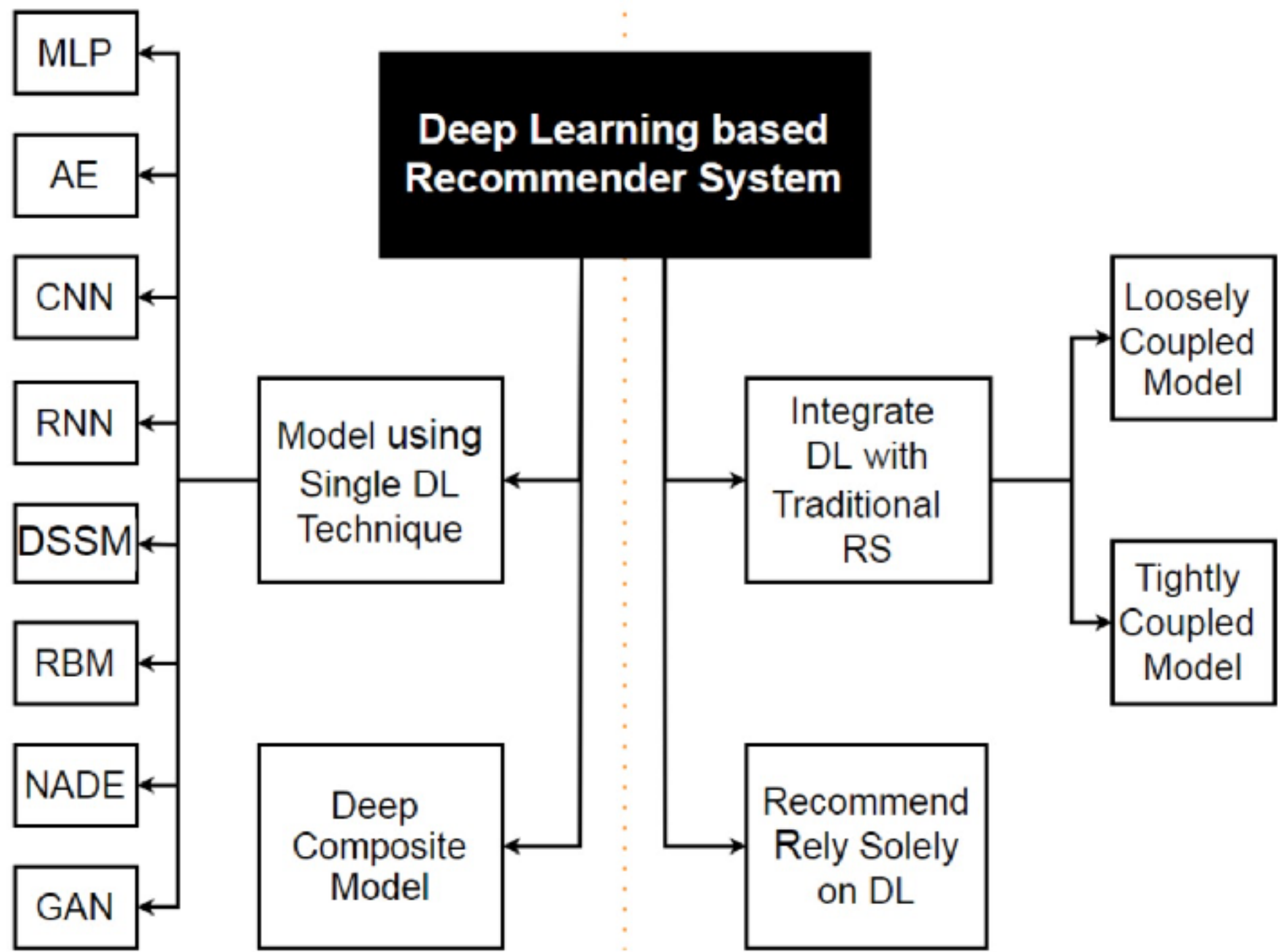
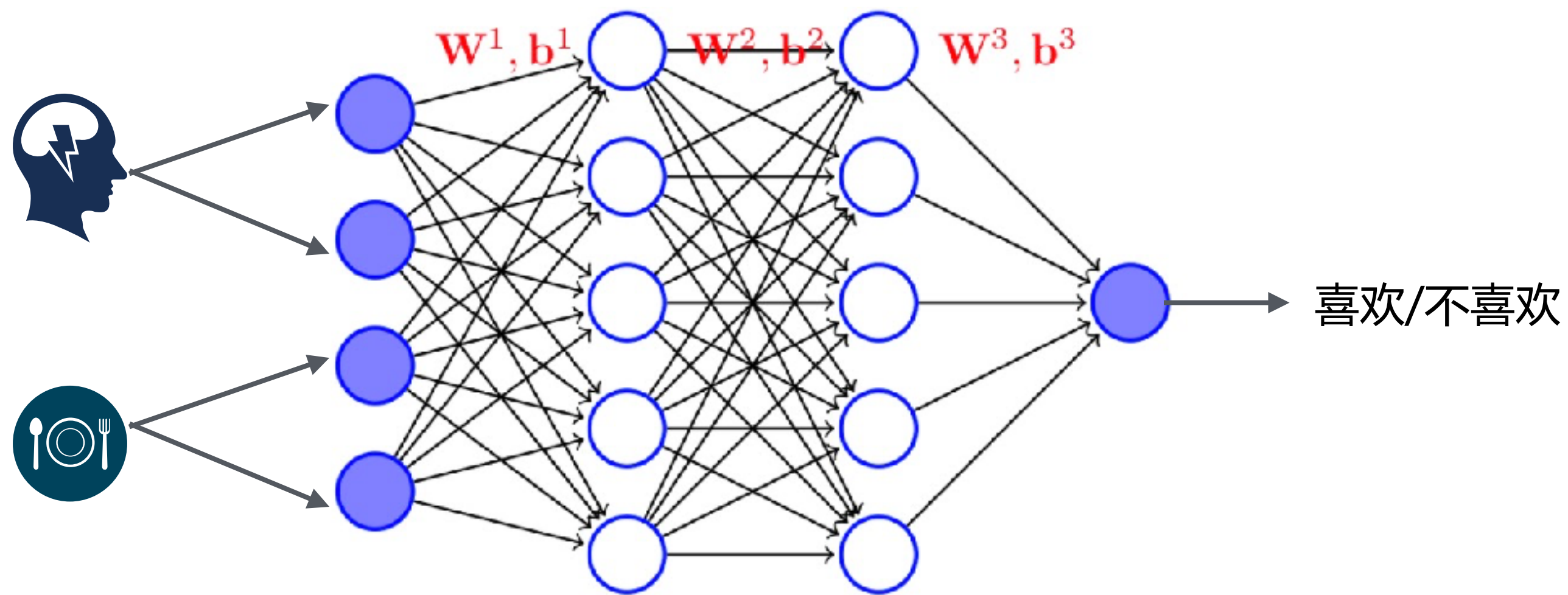
- 产品特点

- 内容源丰富，包括图文、商户、视频、社交关系等
- 静态内容与动态相结合、非单纯信息流

- 深度学习应用

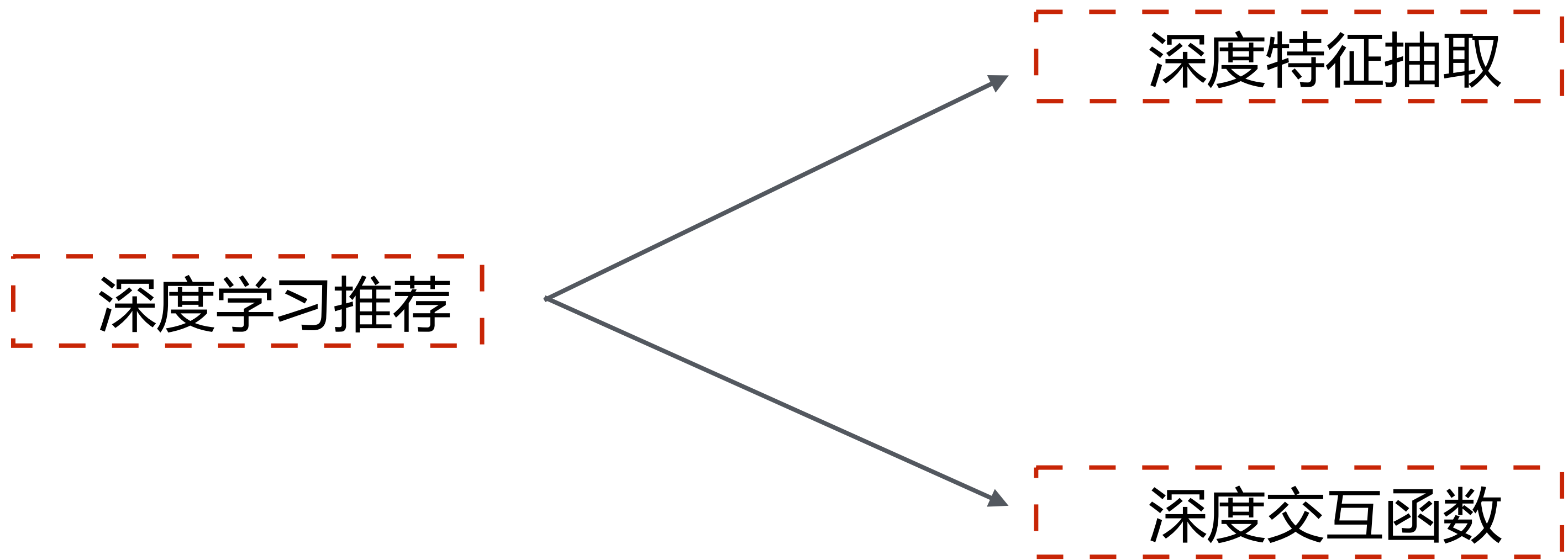
- 召回模型：DSSM、Word2Vec、LSTM
- 排序策略：Wide&Deep、DNN
- 展示层优化：CNN



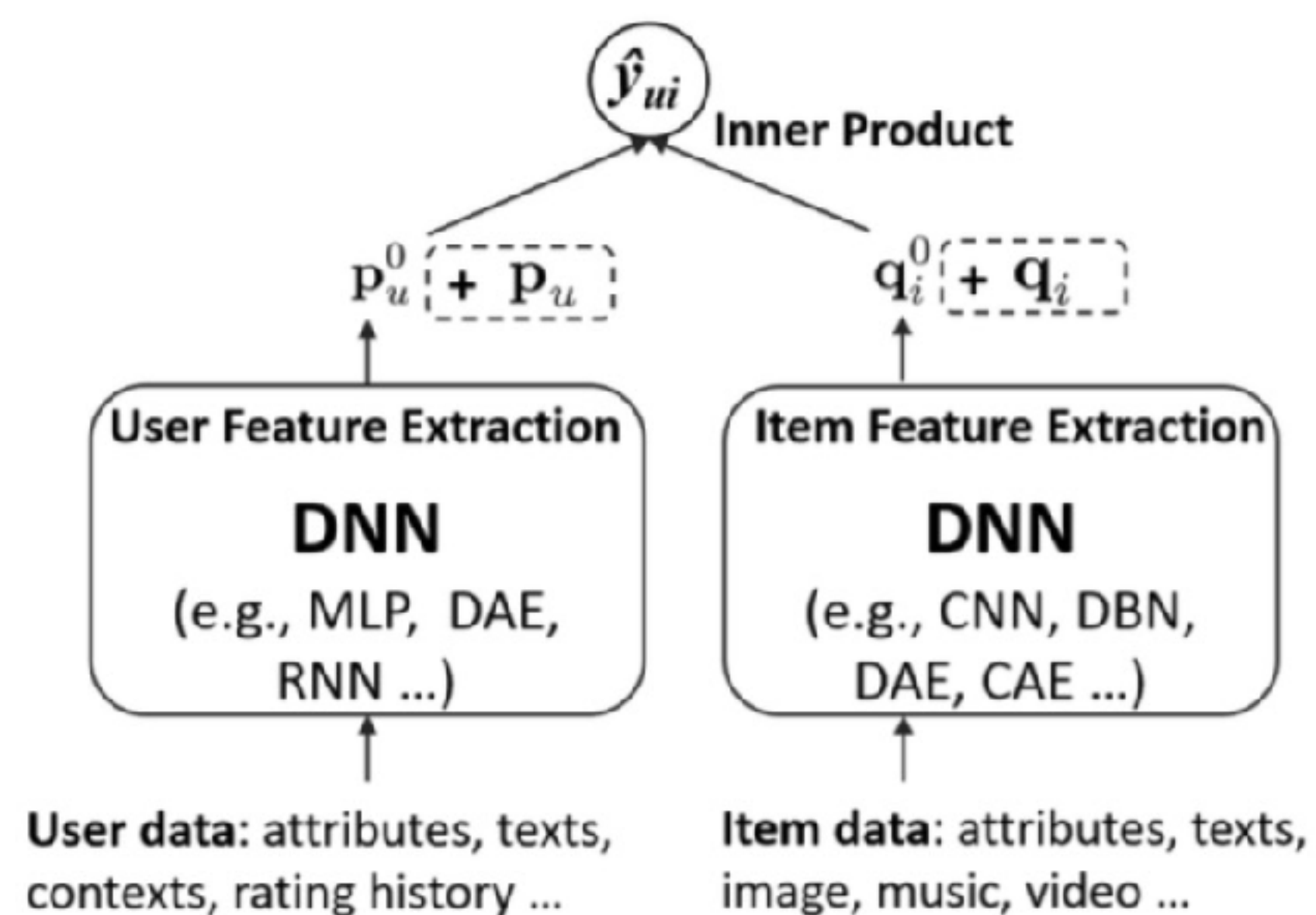


参考文献

[1]Shuai Zhang et. Deep Learning based Recommender System: A Survey and New Perspectives

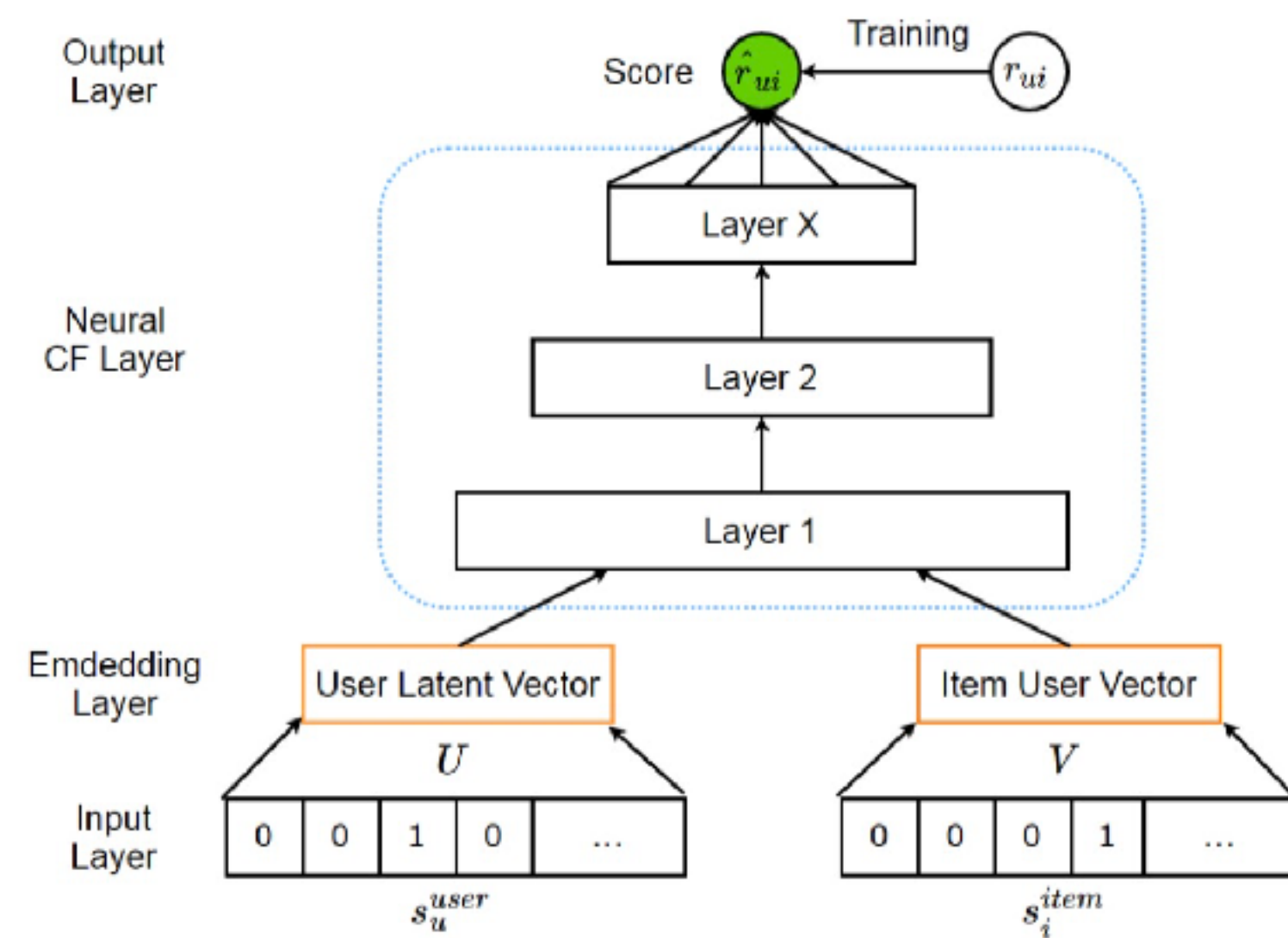


深度特征抽取

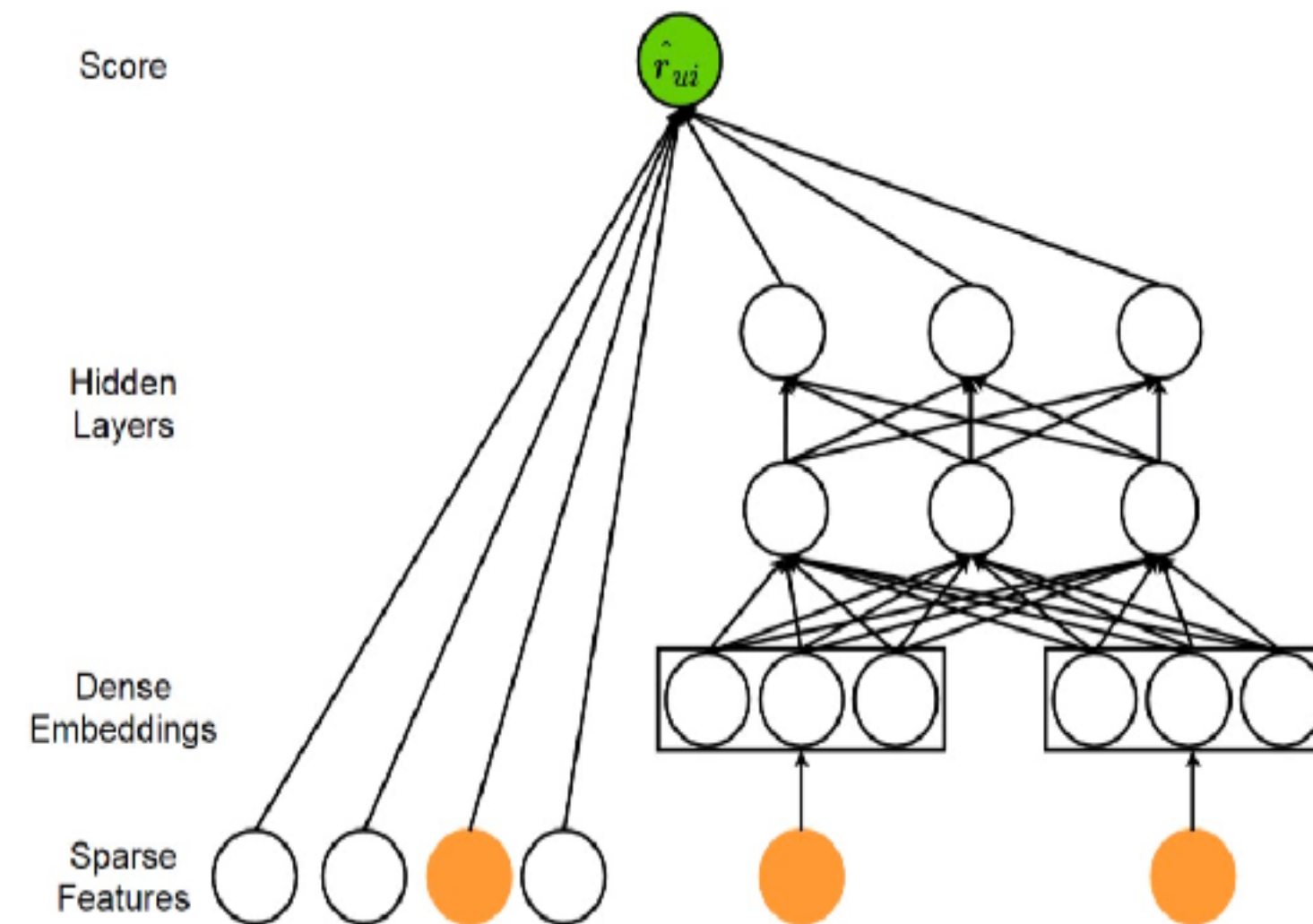


- 基于数据表征的学习，根据Item内容抽取不同的信息
- 采用向量内积作为用户和物品的交互函数

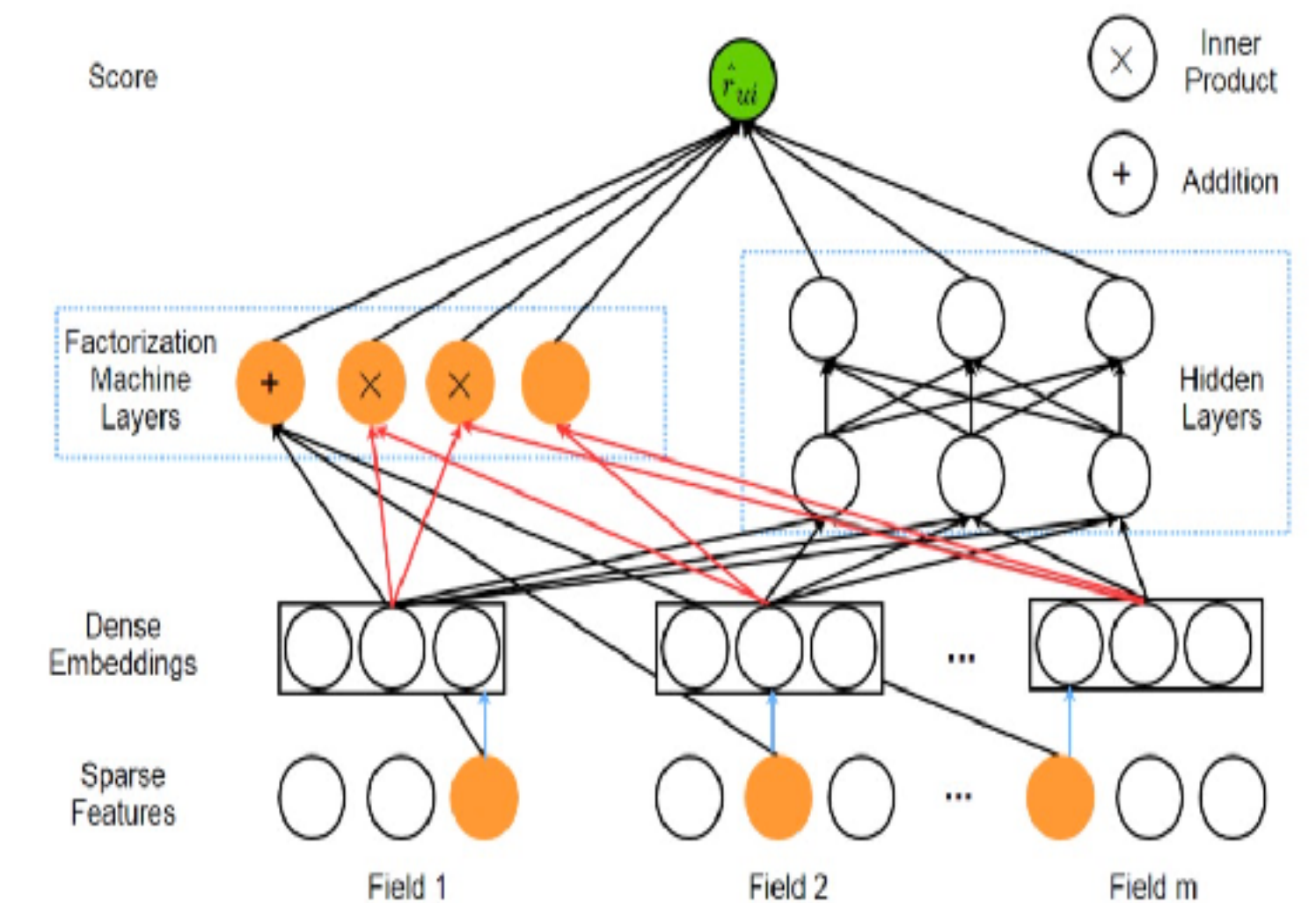
深度交互函数学习



Neural Collaborative Filtering



Wide Deep Learning



Deep FM

- 使用深度网络学习User与Item的交互函数
- 可以在隐含层进行多种组合，建模User与Item之间的非线性关系

- 内容分发

- 项目背景

- 用户模糊决策需求多，通过内容获取并留住高净值用户，提升用户粘性。
 - 为用户推荐感兴趣的内容

- 主要挑战

- 内容标签体系不完备、准确度不高
 - 内容来源多、质量参差不齐
 - 内容热度分化严重、分发效率低

- 解决问题

- 使用基于内容推荐策略解决冷启动问题
 - 考虑文章标题、正文等特征，未来融入图片特征
 - 融入用户行为进行建模，使用DSSM模型



- 内容分发-DSSM原理

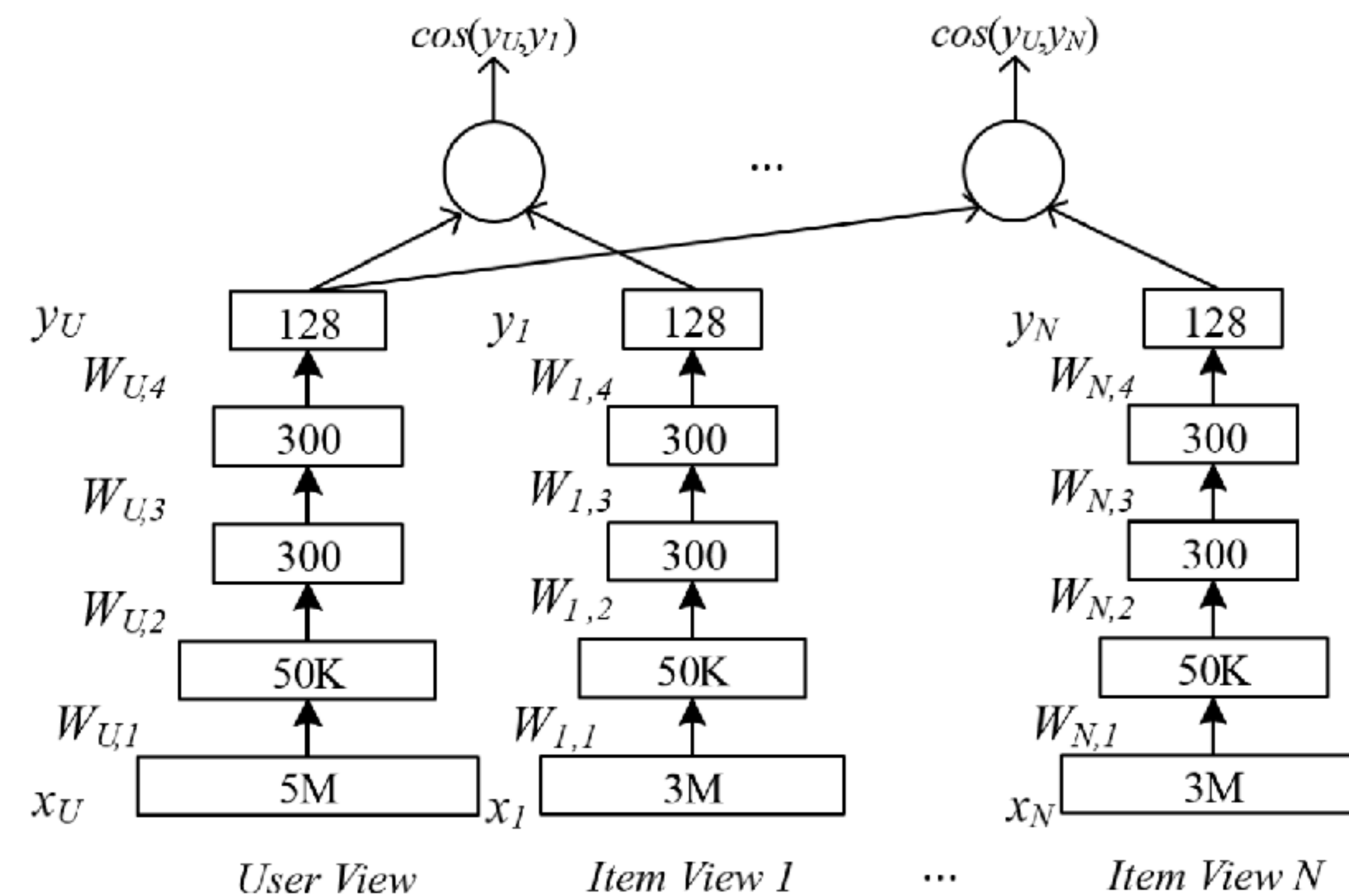
- 使用DNN模型将用户和内容转化为统一的语义空间

- 计算步骤

- 根据用户行为数据，进行过滤和采样，抽取点击对
- 对用户和内容进行Embedding
- 相似度计算，进行内容推荐

- 核心挑战

- 数据的清洗与点击对构造
- 行业相关的内容实体抽取



- **DSSM实现环境**

- **实现方式**

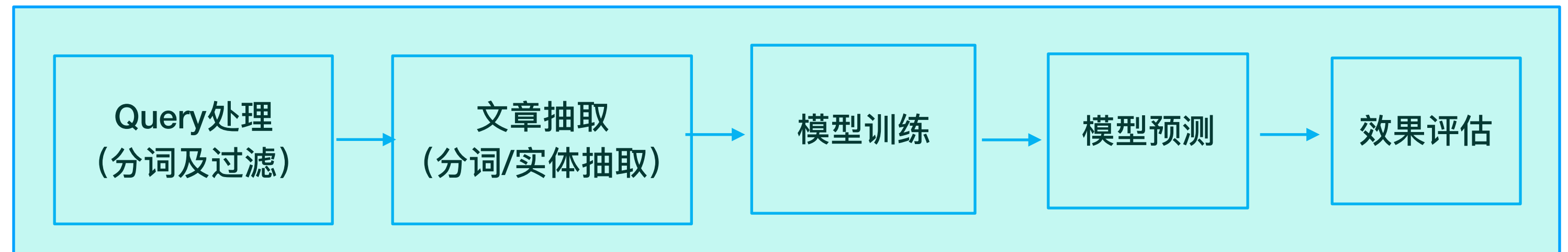
- 微软CNTK

- **样本对选取**

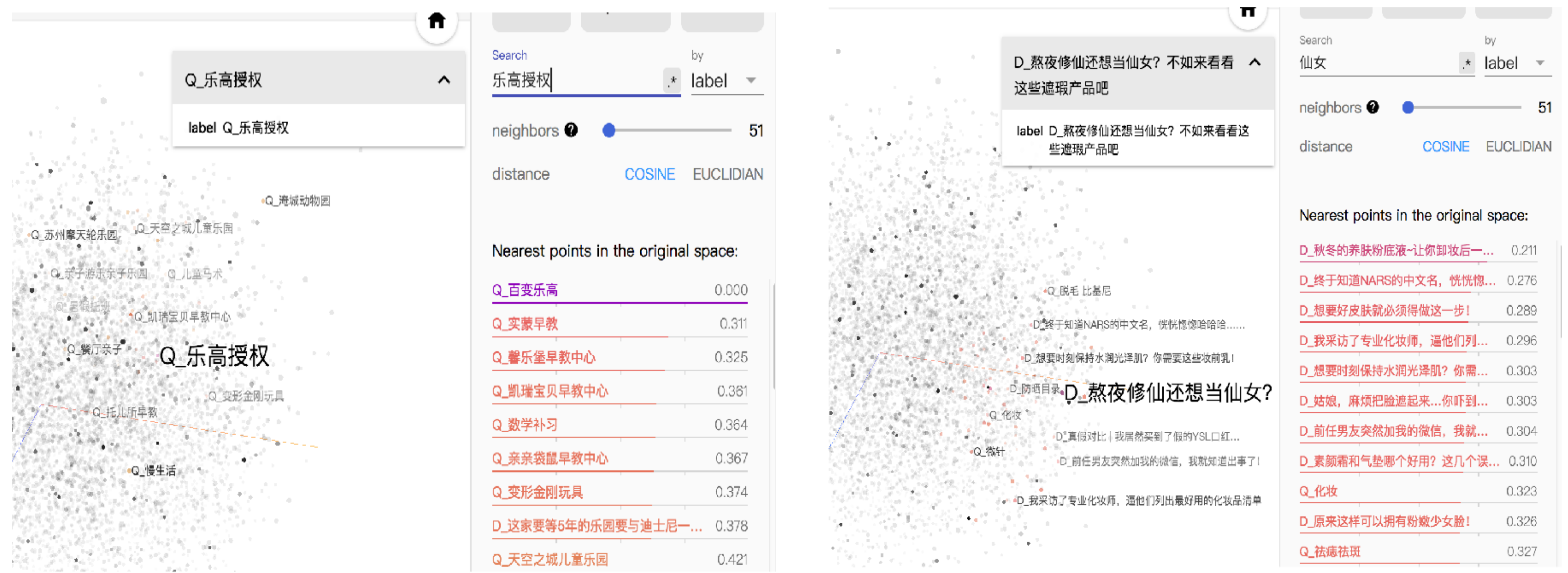
- 搜索query与点击的文章
 - 点击POI名称作为query及点击文章

- **训练**

- 1000w样本、30次迭代，2小时
 - CrossEntropy作为损失函数

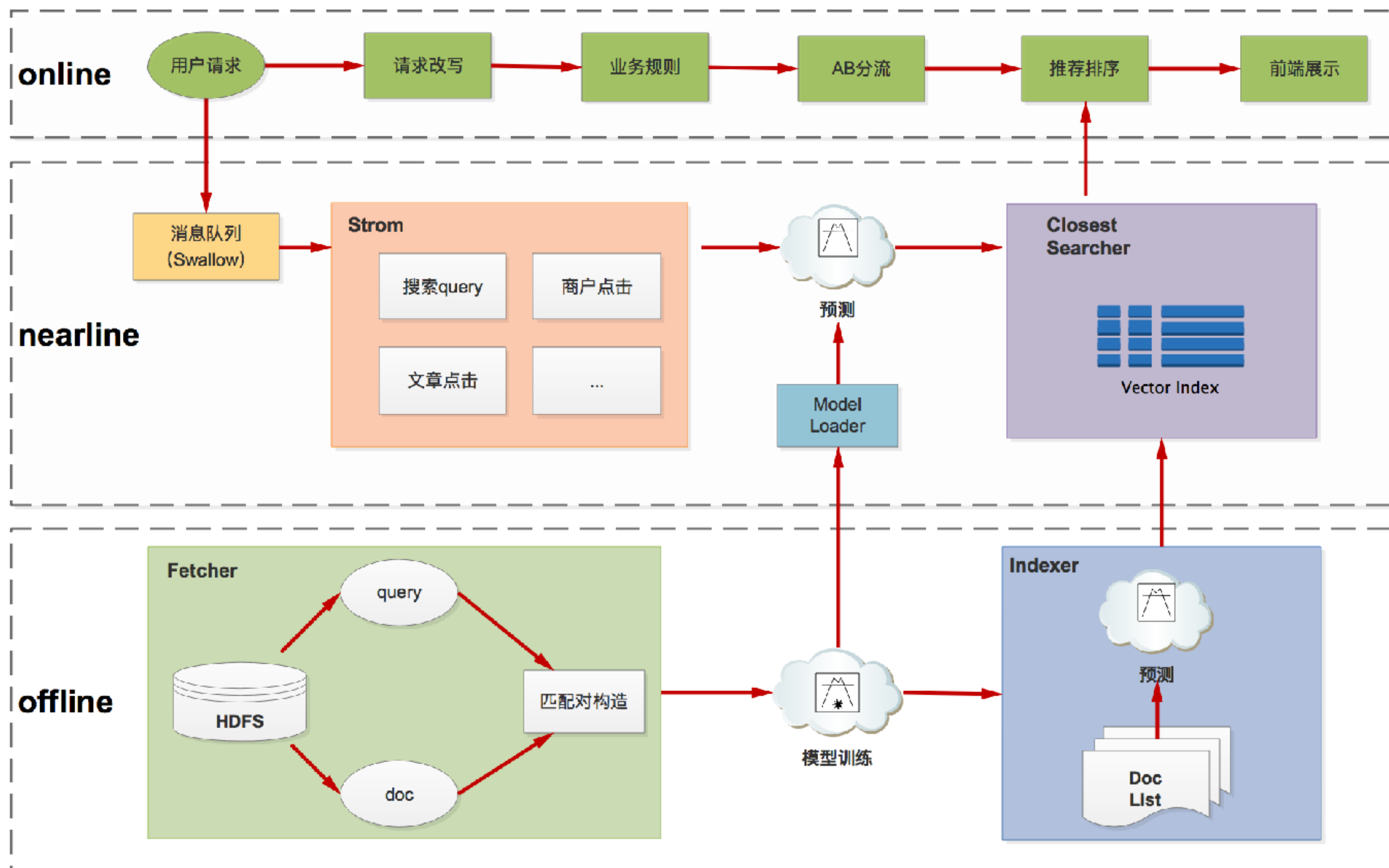


- DSSM实验效果

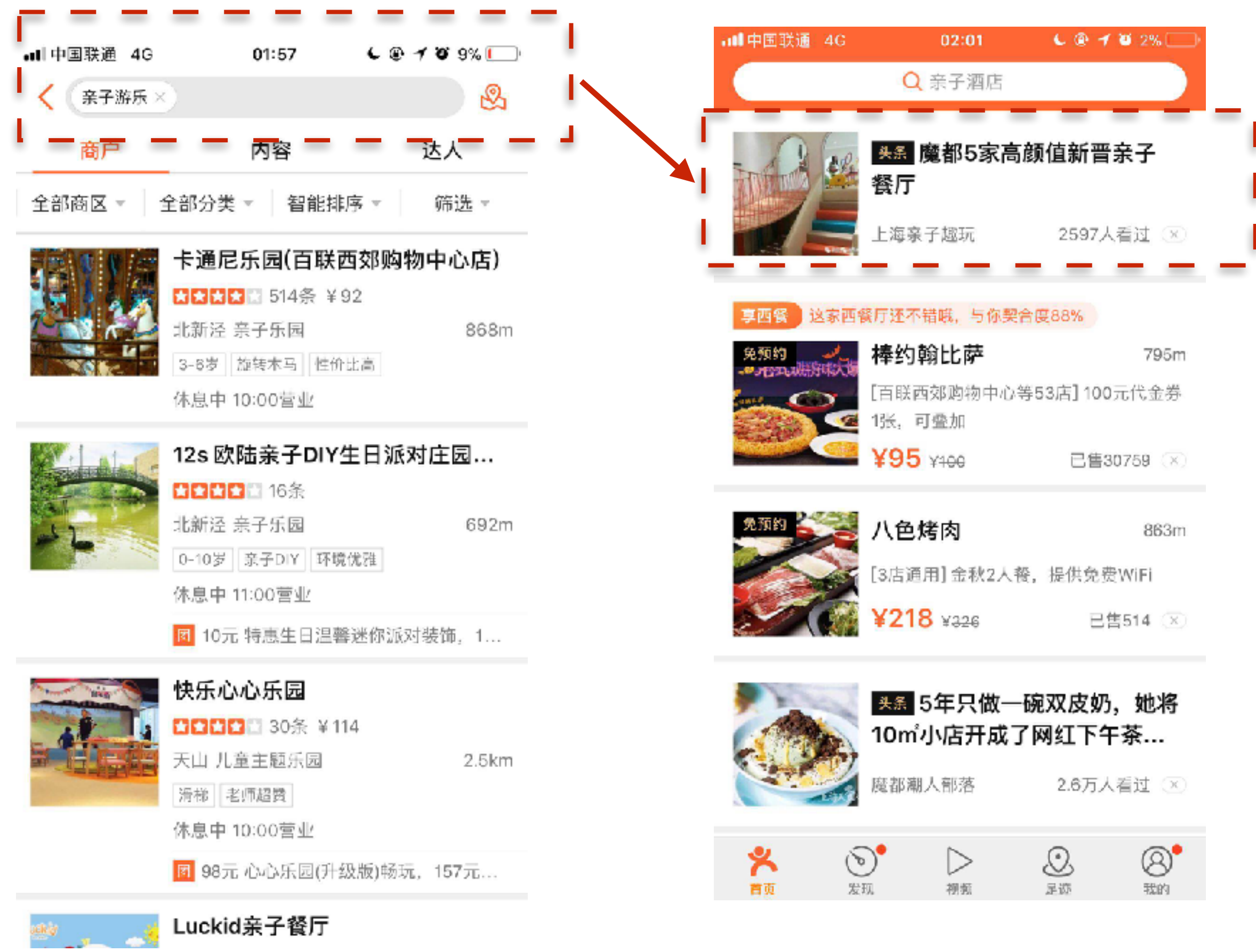


语义相似效果

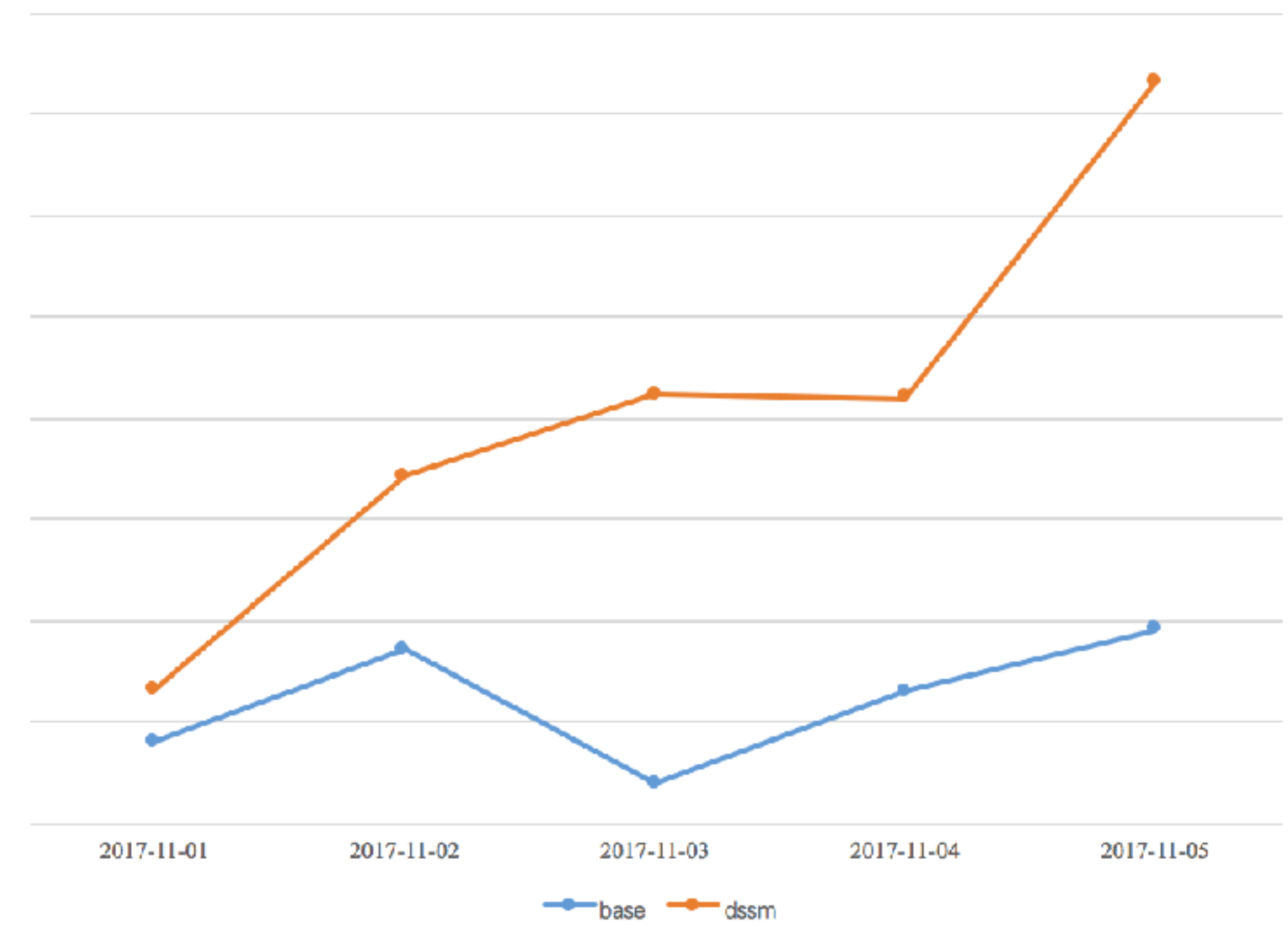
- DSSM线上流程



• DSSM算法效果



DSSM线上效果



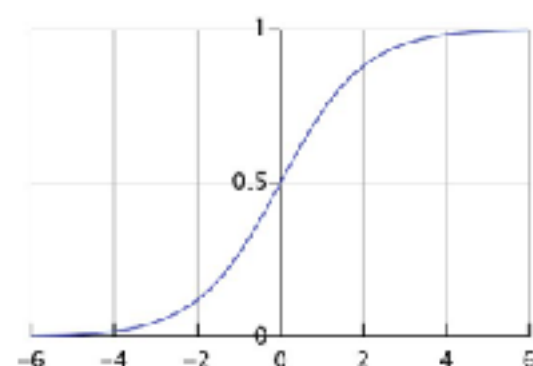
线上效果提升

对推荐算法产生的候选集，为什么要**深度学习排序**？



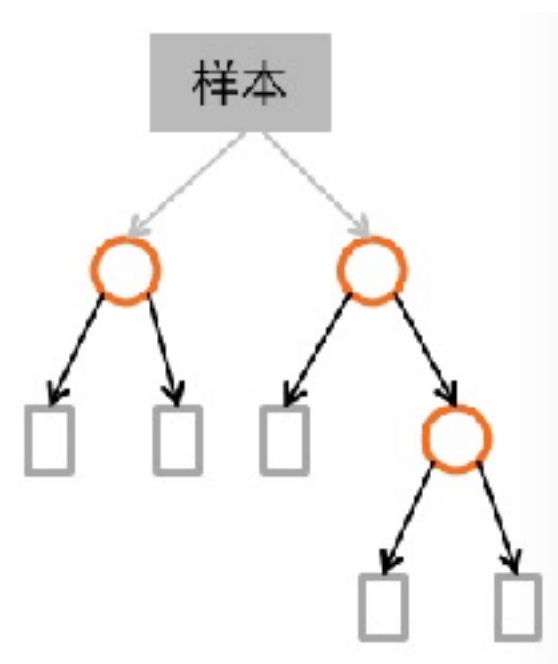
传统排序算法对比：

线性



- ✓ 记忆能力强，泛化能力弱，对特征工程要求高
- ✓ 需要人工组合高阶特征

树模型



- ✓ 非线性模型可以探索拟合特征向量空间，泛化能力增强
- ✓ 可以显示出特征的重要度，对单特征值误差容忍度高
- ✓ 无法利用大规模稀疏特征

FM/FFM

$$y(\mathbf{x}) = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n w_{ij} x_i x_j$$

- ✓ 泛化能力强，可以组合二阶特征
- ✓ 高阶特征组合难

技术角度

- 特征多且特征工程耗时
- 稀疏特征利用的少
- 模型泛化能力弱

推荐排序

- Wide模型记忆好
- Deep模型泛化好
- 可以增加交叉特征

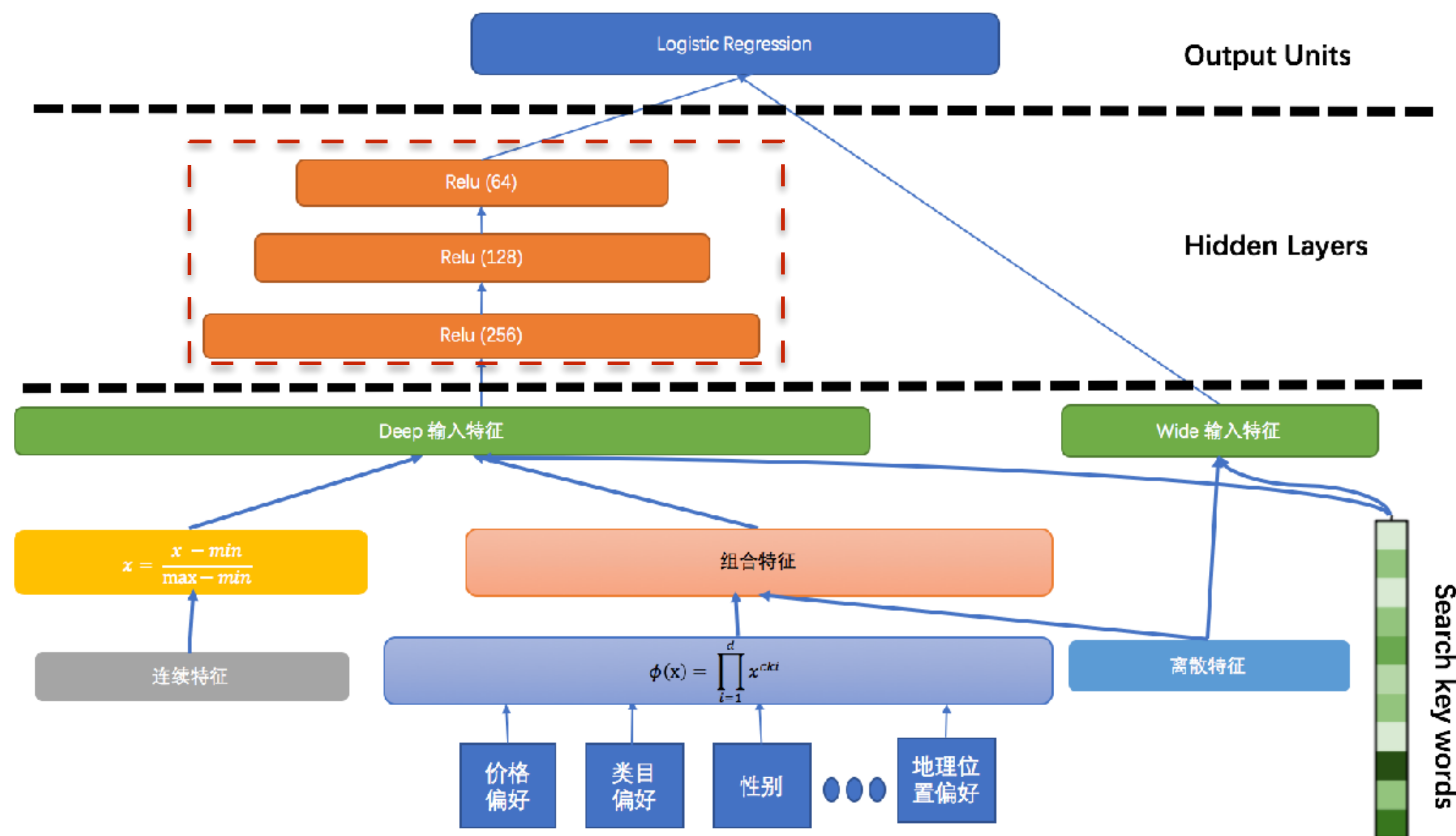
产品体验方角度

- 能够适应场景的变化
- 需要新颖性

- 不能兴趣过载
- 业务效果有提升

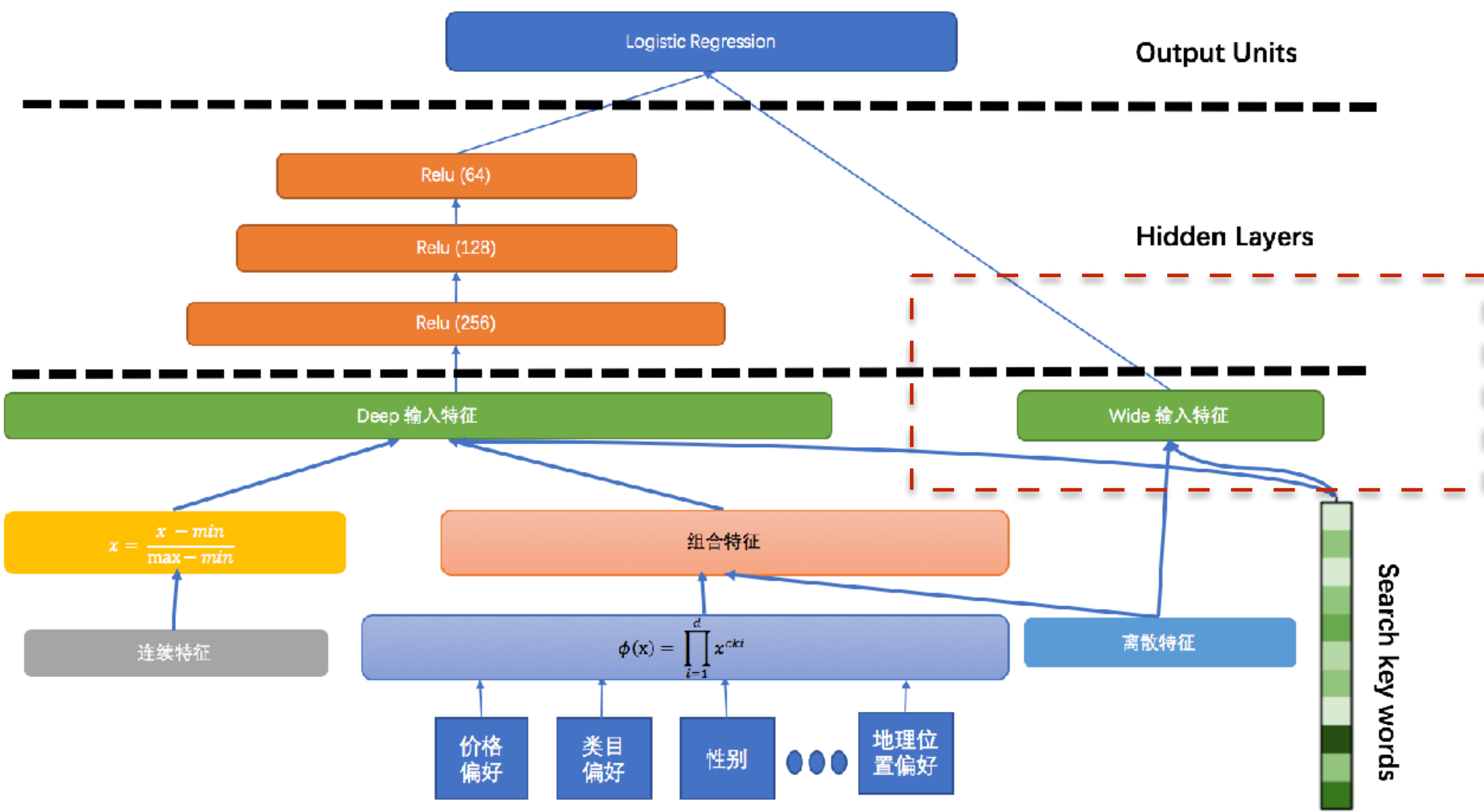
Wide&Deep排序

Wide&Deep排序算法：



- Deep网络负责泛化：
离散特征分域 Embedding
连续特征按照累计分布函数归一化

Wide&Deep排序算法：



- Wide网络负责记忆：
浅层网络需要特征设计
CTR等统计特征适合在浅层网络

特征构建



特征处理

特征处理

- 归一化
- 标准化
- 离散化
 - 等频、等值
- 数据变换
 - 多项式、指数、对数

特征组合

- 人工组合
- 模型组合
 - GBDT
 - FFM

特征编码

- One Hot Encoding
- Word2Vec

Wide&Deep实现

- 训练工具
 - Keras
- 训练数据
 - 7000W训练样本、3000W测试样本
 - Batch Size为50000, Epoch为20
- 训练参数
 - 激活函数: ReLU (前三层)、Sigmoid (最后一层)
 - 损失函数: Cross Entropy
 - 优化器: Adam
- 线上部署
 - DeepLearning4J
 - ND4J库进行矩阵运算加速
 - 线上预测速度快



DEEPLARNING4J

Wide&Deep线下实验效果：

算法模型	AUC
基础模型	0.6885
Deep learning	0.6998
Wide &Deep	0.7065
Wide&Deep (Cross Feature)	0.7181

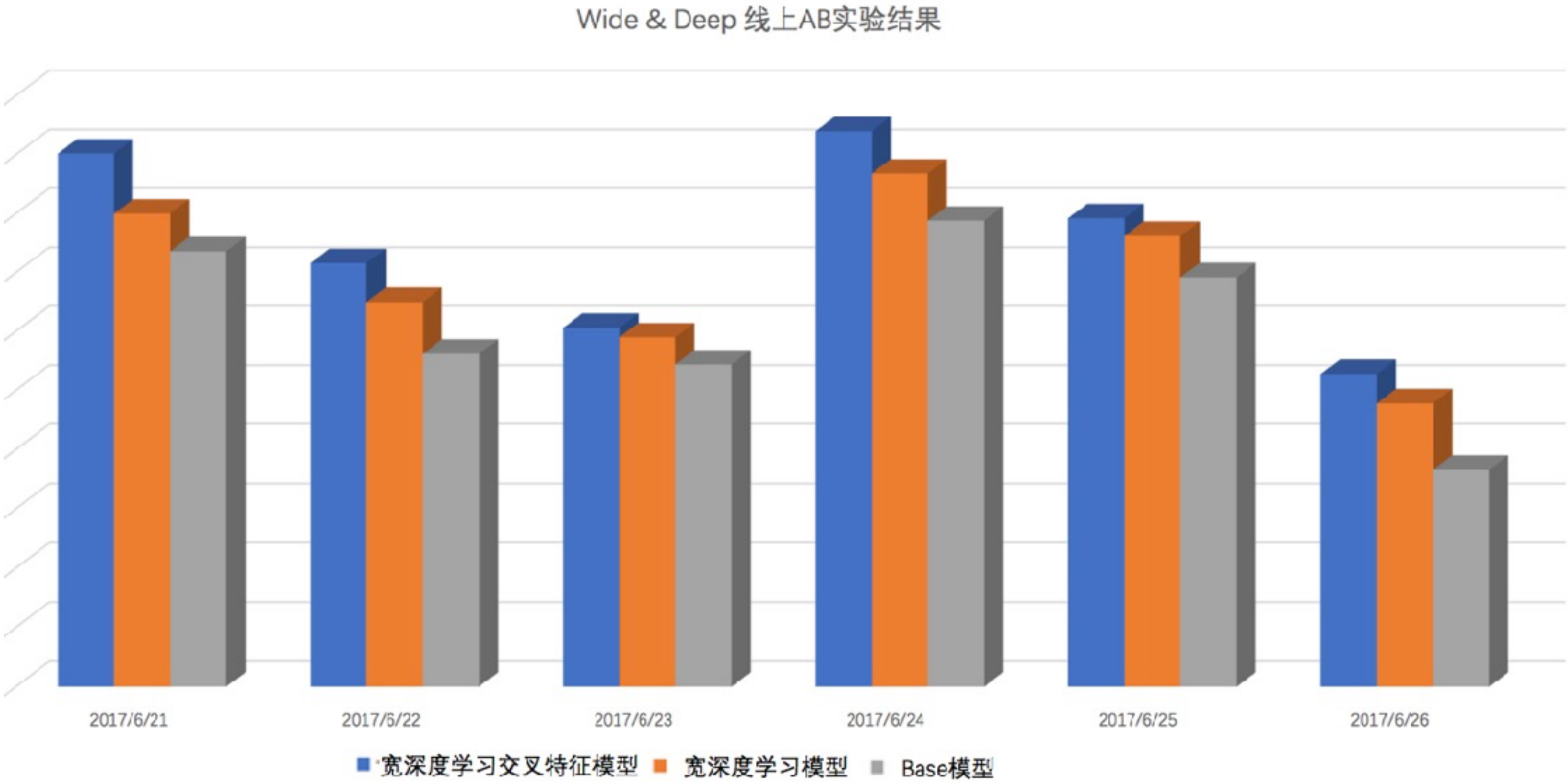
- 结论：Wide&Deep的离线AUC有较为明显提升，增加交叉特征，可以提升离线效果。

Wide&Deep线下实验效果：

隐层宽度	AUC
512 ReLU	0.7221
256ReLU	0.7181
256ReLU->128ReLU->64ReLU	0.7166

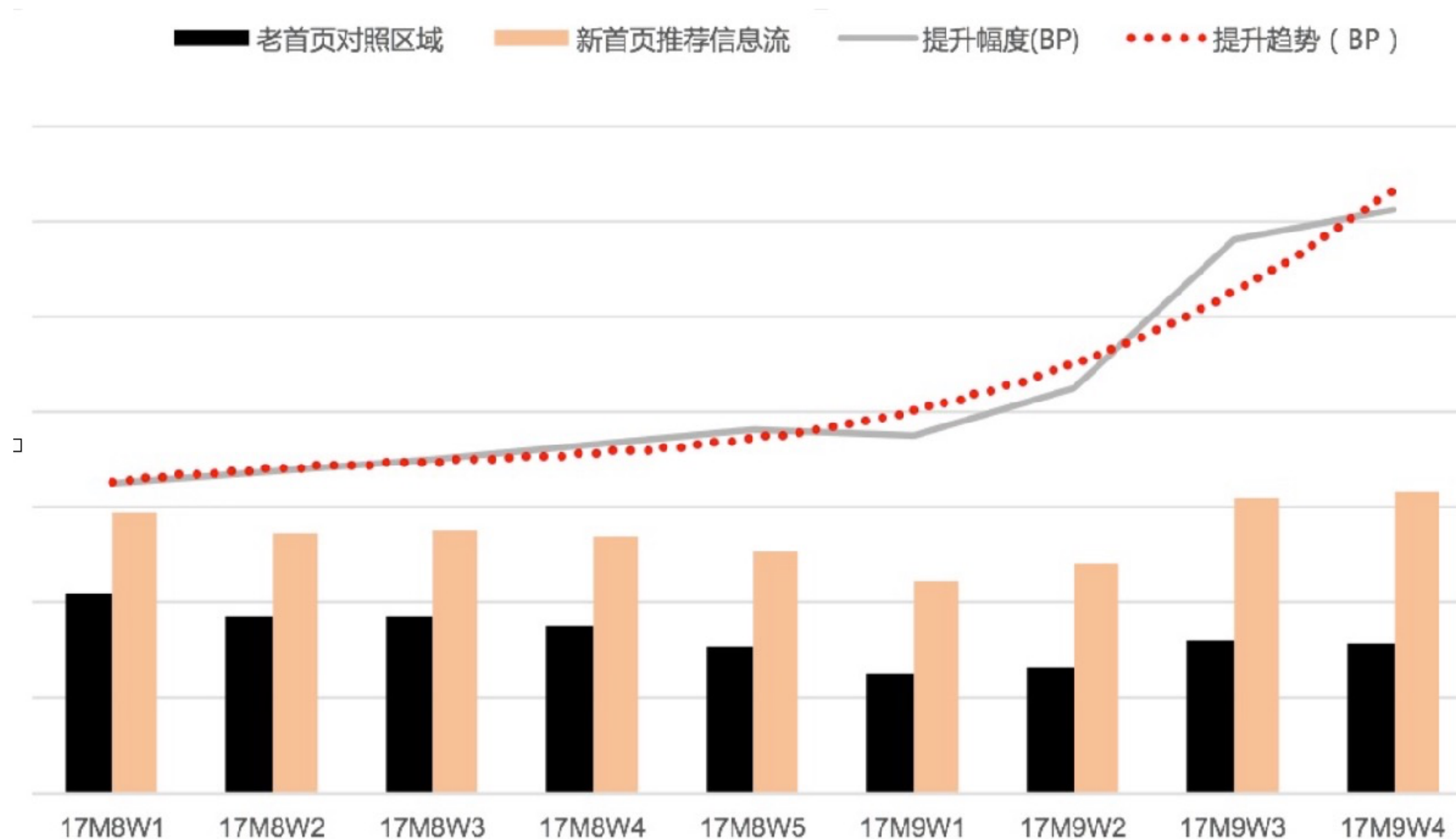
- 结论：随着隐含层的宽度增加，离线评估效果也有相应提升，考虑到线上服务性能，目前使用的层宽还未做进一步扩展。

Wide&Deep排序算法线上效果：

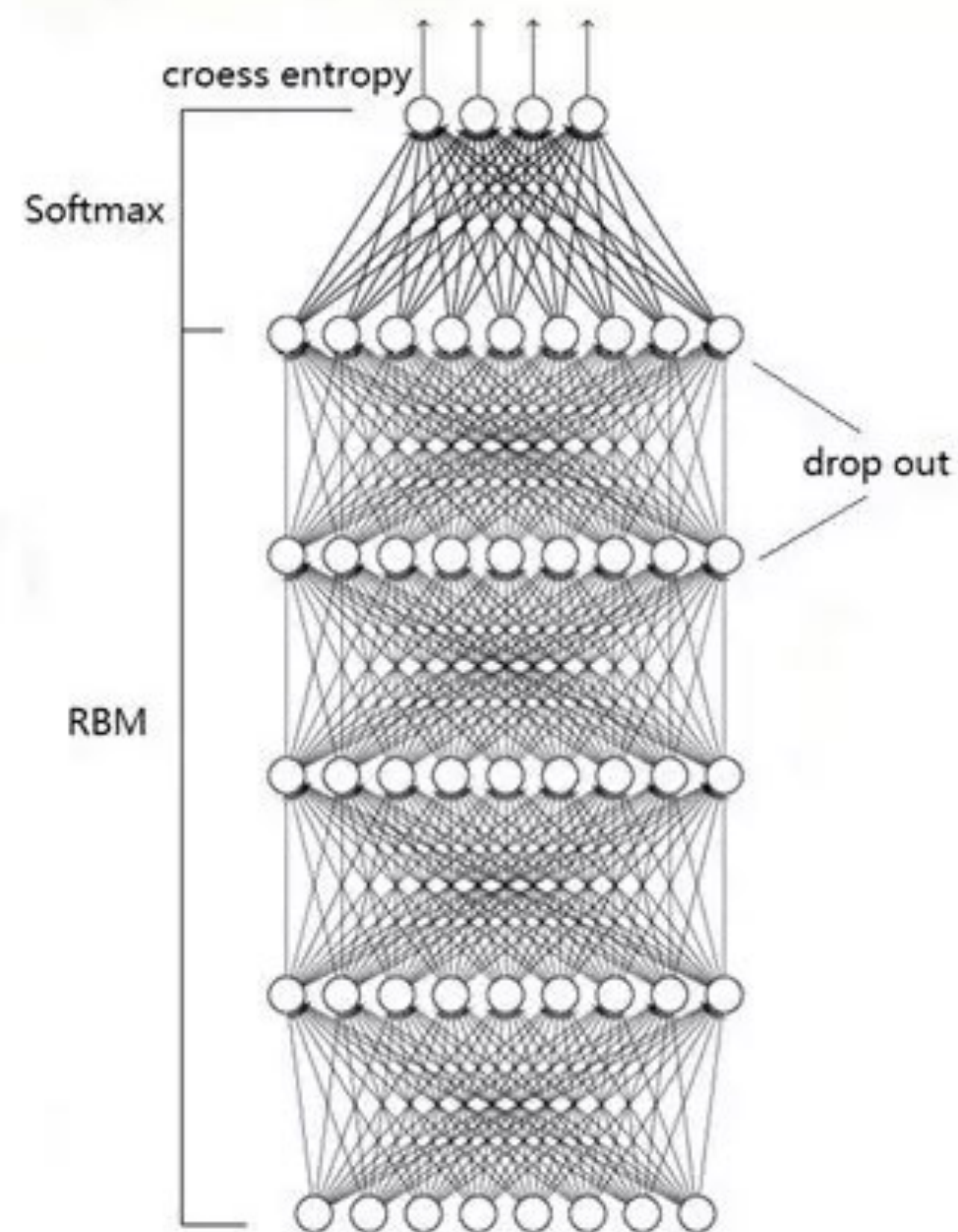


- 结论：使用Wide&Deep模型，线上的核心指标有明显提升，目前已在线上全面应用。

Technology Bridge Business



- 结论：使用Wide&Deep排序等新模型进行业务迭代后，首页Feeds AB测试的效果逐步提升



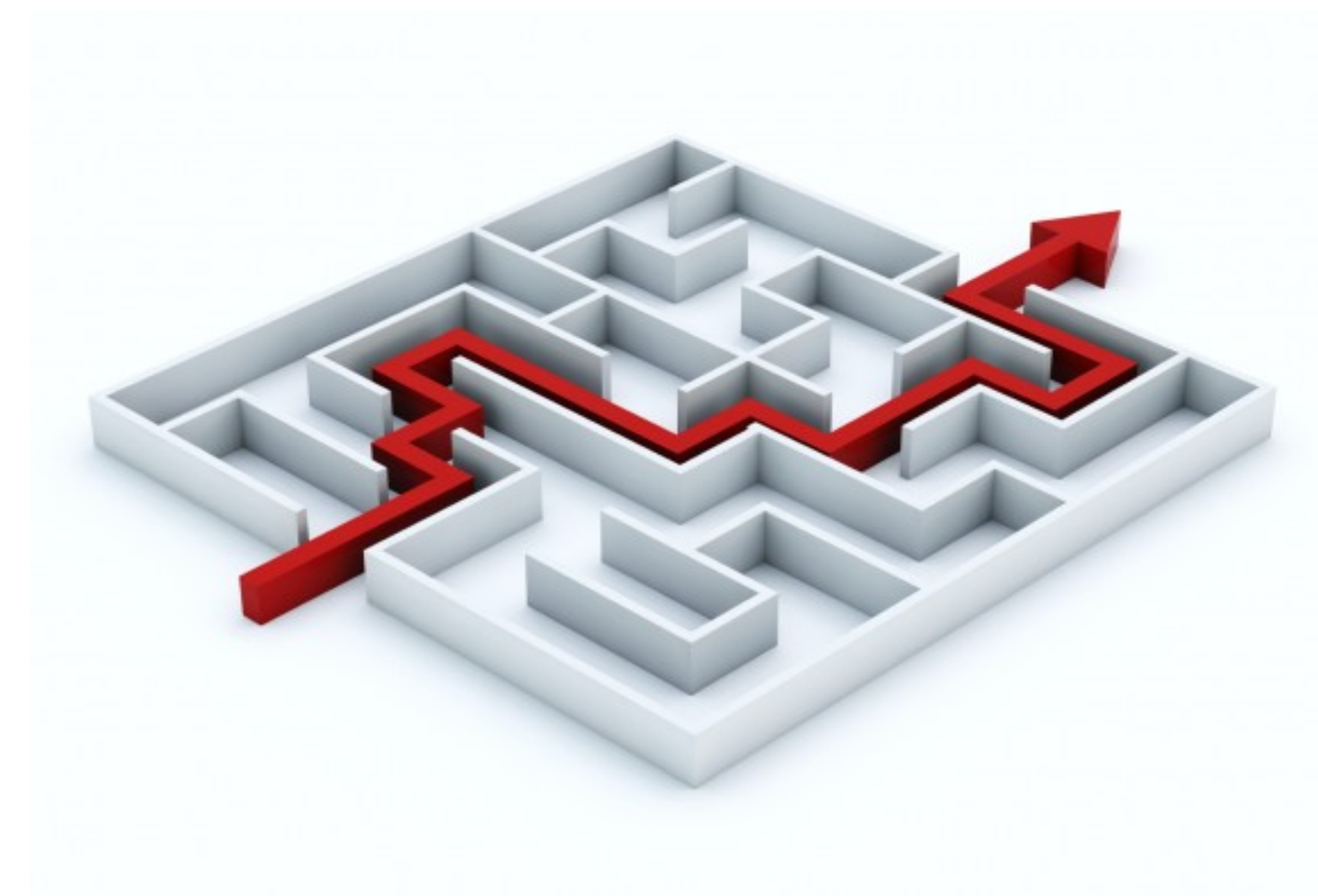
- 深度学习的优点

- 泛化能力强，更多的丰富的特征交叉
- 更多的非线性关系及可调参空间

- 深度学习的不足

- 训练复杂，模型训练和预测耗时长
- 受制于线上性能，目前网络复杂度有限
- 结果黑盒，稀疏特征需要Embedding

- 模型优化
 - 更宽、更深、更复杂的网络结构，例如DeepFM模型
 - 近线层深度召回模型开发，例如考虑行为序列的LSTM模型
 - 高吞吐、低延时的在线服务性能优化
 - 深度增强学习
- 特征融入
 - CNN挖掘图像特征
 - Doc2Vec特征挖掘



Q&A



