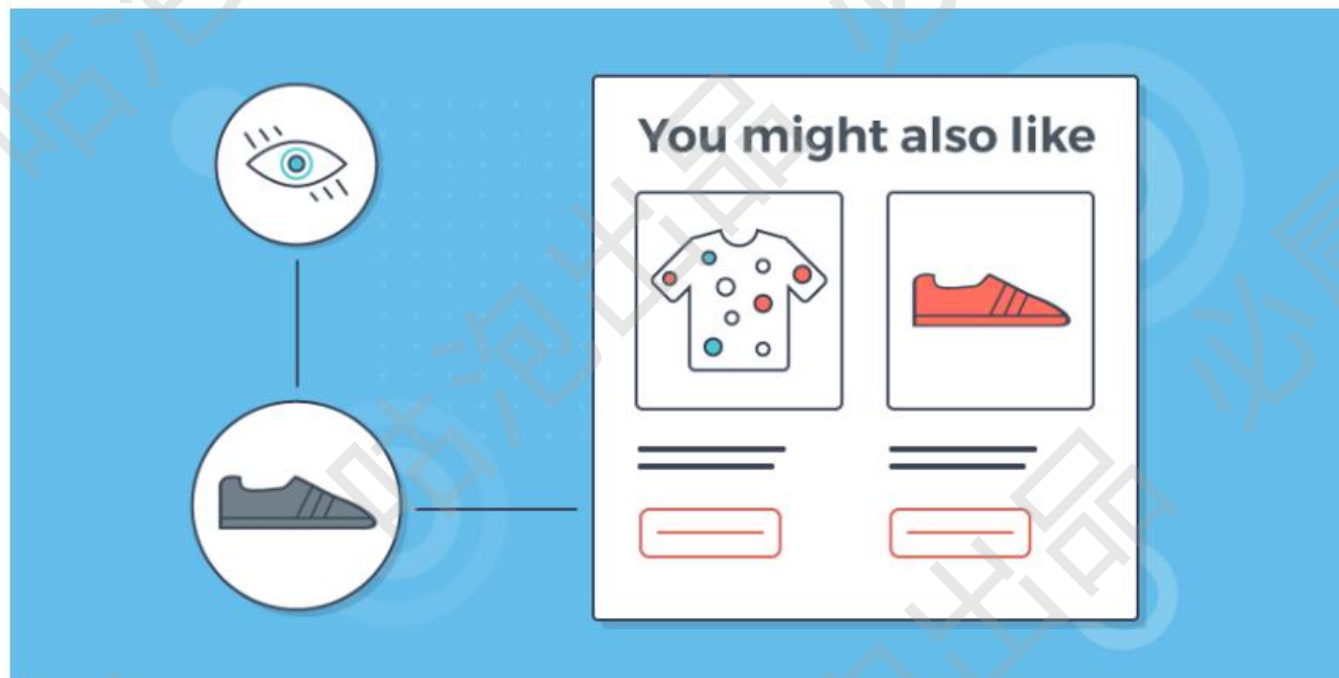


推荐系统

✓ 什么是推荐系统呢？

✎ 高度总结：来了就别想走了，抖音一刷一整天，微博越刷越过瘾



推荐系统

✓ 好产品都推荐的好

✎ 大数据时代也带来了一些问题，信息太多了，到底哪个是我需要的？

✎ 话题很多，感兴趣的却很少，这就需要推荐系统来登场了

✎ 抖音越刷越是自己喜欢的东西（导致不敢让别人别我抖音了）

✎ 京东越买越想买（导致不敢乱搜索了，只是个段子。。。）



推荐系统

- ✓ 推荐系统无处不在
- 📎 你可能感兴趣的，你可能喜欢的
- 📎 越是大厂，推荐做的越好
- 📎 千人千面，所有人的推荐都不一样
- 📎 用户的行为决定了推荐的内容



推荐系统

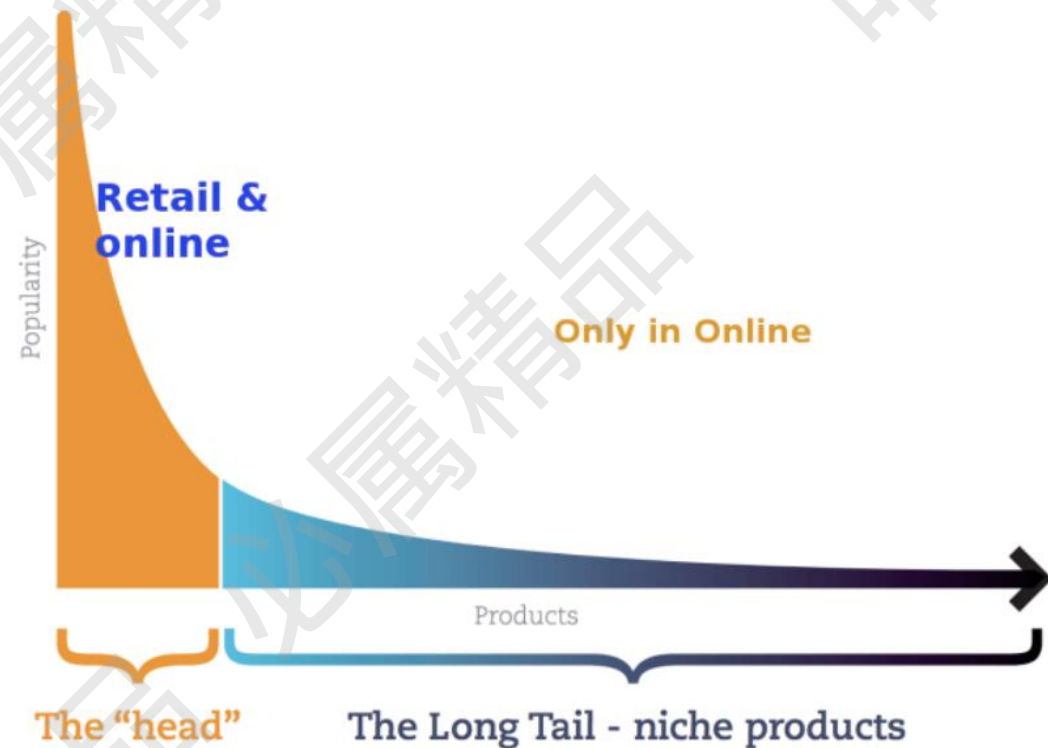
✓ 为什么需要推荐系统

✎ 卖的好的就那几种商品，其他的不要了？

✎ 80%的销售来自20%的热门商品

✎ 要想提高整体收益，得把长尾商品推出去

✎ 通过用户行为来进行个性化推荐





推荐系统

✓ 推荐系统发展

✎ 亚马逊1998年就开始用了，只不过那时候还是简单的协同过滤

✎ 2006年Netflix（在线视频）竞赛一炮走红

✎ 2015年开始深度学习崛起，推荐也随之改变

✎ 今天已经百花齐放，各大论文层出不穷



to rest over sink's drain
• Easy wash with warm soapy water to keep the mirror finish shine impeccable over time. Dishwasher safe.
• See more product details
Compare with similar items
New (5) from \$9.99 & FREE shipping on orders over \$25.00 shipped by Amazon. Details

Father's Day
Up to 30% off tools
See more +

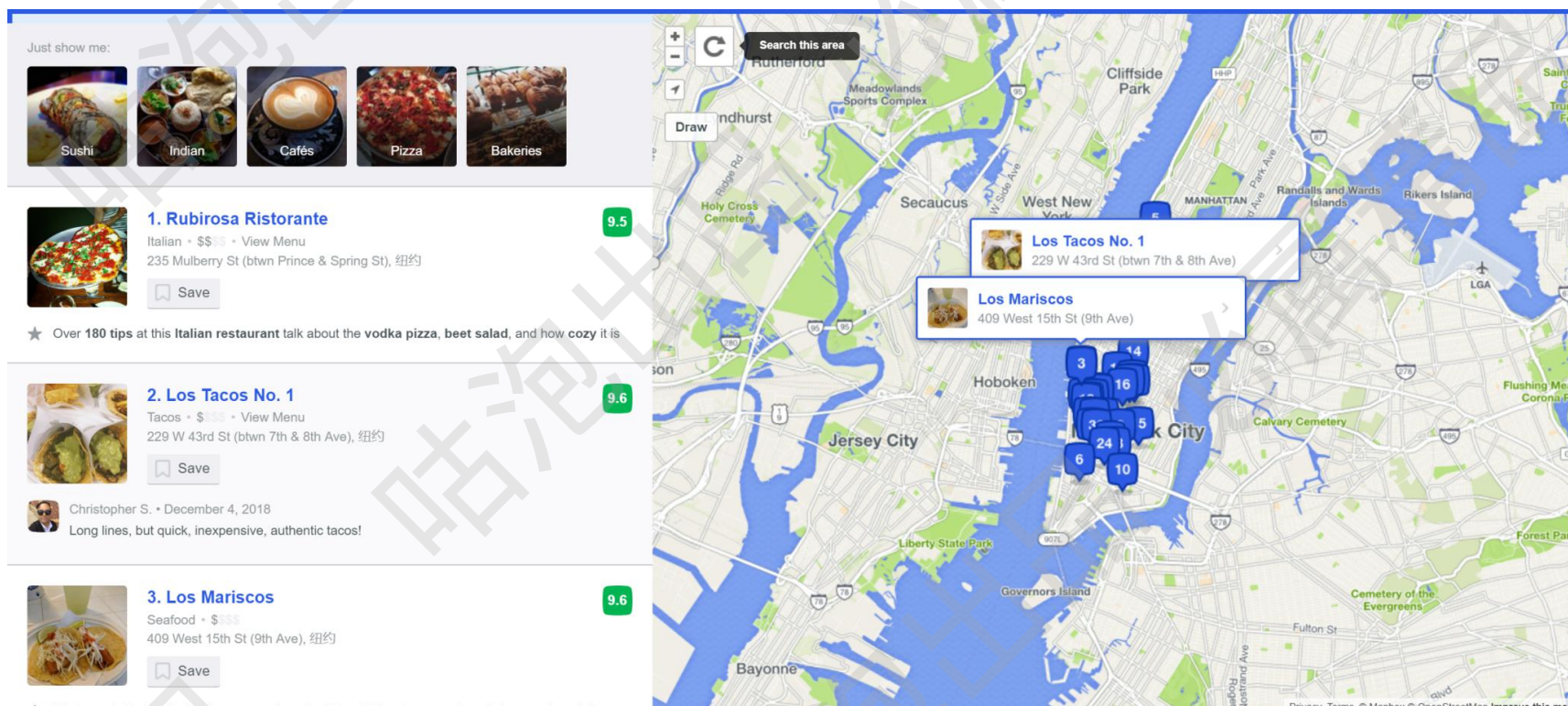
Father's Day gifts in Home Gift Guide
Discover gifts that will delight dad. From tools & equipment to kitchen gadgets and grills, find inspiration for Father's Day gifting. Sponsored by Bosch Shop now



推荐系统

✓ 推荐系统应用领域

📌 广告都是钱，地图也得回回本，到哪了都有一堆推荐！



推荐系统

✓ 推荐系统应用领域

✎ 不仅仅是推荐产品，还打辅助，帮助完成各种决策

✎ 为什么她会成为可能认识的人？

✎ 图模型能做的事情实在太多啦，万物互联

✎ 知识图谱也是做好产品必备技能



推荐系统

✓ 推荐系统应用领域

✎ 个性化推荐，优化用户体验，海量数据中快速定位，精准营销



推荐系统

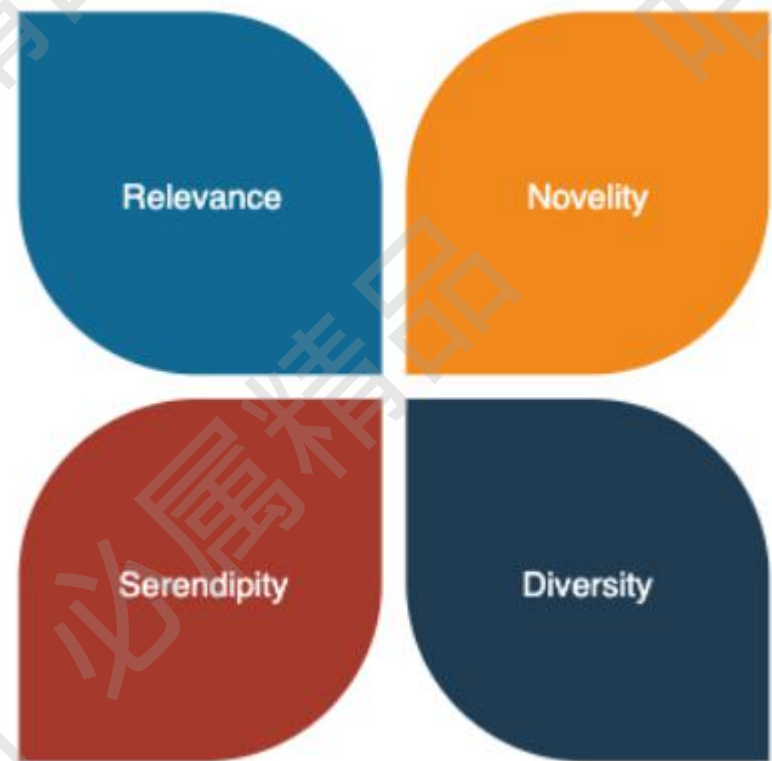
✓ 推荐系统的目标

✎ Relevance: 推荐的东西起码得相关才行

✎ Novelty: 新颖的才好, 推的得是人家没有的

✎ Serendipity: 跟处对象一样一样的, 机缘

✎ Diversity: 多样性, 换着花样玩才好



推荐系统

✓ 如何评价推荐系统效果

✎ 其实最主要的还是用户满意度，各种数学公式只是辅助判断

✎ 常规的计算损失：
$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{T}|} \sum_{(u,i) \in \mathcal{T}} (\hat{r}_{ui} - r_{ui})^2}$$
 (T是测试集)

✎ TopK推荐：
$$\text{precision@K} = \frac{|Rel_u \cap Rec_u|}{|Rec_u|}$$
 (用户相关商品集与推荐商品集的交集)

✎ 覆盖率：挖掘长尾；多样性：覆盖不同领域；实时性：刚买了房子赶紧推装修

推荐系统

✓ 推荐系统里的常用词

✎ Item：商品，例如要从拼多多买的9.9包邮的拖鞋

✎ Embedding：隐向量，例如对用户商品评分矩阵进行分解

✎ 召回：粗略计算要返回结果，例如先从100W商品中取比较可能的100个

✎ 打分：要排名得有一个统一的标准；重排：最终结果排序

推荐系统

✓ 推荐系统经典流程

✎ 离线+近线+在线（召回+粗排+精排）


✎ 离线通常跑较大的模型与算法，先得到当前数据的大致结果，一定时间更新一次

✎ 粗排通常会跟着用户走，用户做了什么事，推荐结果也会随之更新

✎ 在线模块需要根据业务规则来返回最终呈现结果

推荐系统

- ✓ 推荐系统难点与挑战
 - ✎ 需要更广泛的收集用户标签并画像
 - ✎ 人是善变的，随着时间的推移，兴趣也会改变
 - ✎ 根据固定画像数据，推荐结果可不能固定不变
 - ✎ 特征工程如何构建一直是一个大难题



✓ 推荐系统难点与挑战

- ✎ 需要更广泛的收集用户标签并画像
- ✎ 人是善变的，随着时间的推移，兴趣也会改变
- ✎ 根据固定画像数据，推荐结果可不能固定不变
- ✎ 特征工程如何构建一直是一个大难题



推荐系统

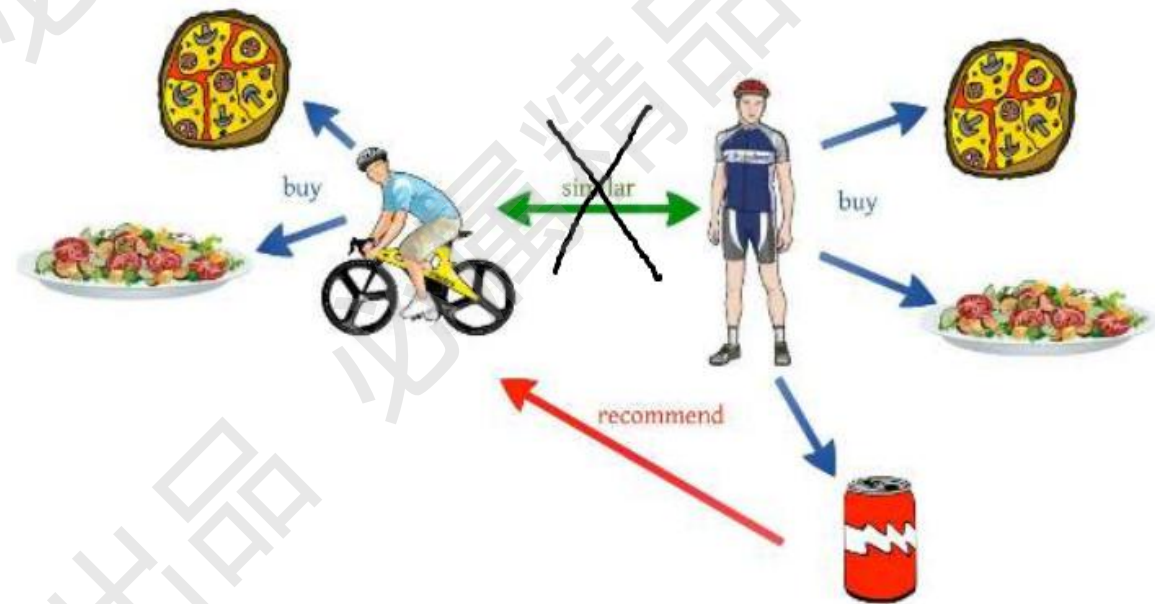
✓ 推荐系统难点与挑战

✎ 冷启动怎么办，包括用户冷启动与商品冷启动

✎ 新用户来了，不知道他啥样怎么办

✎ 商品倒是好办，属性相对固定

✎ 解决方法比较多，例如直接推荐销冠

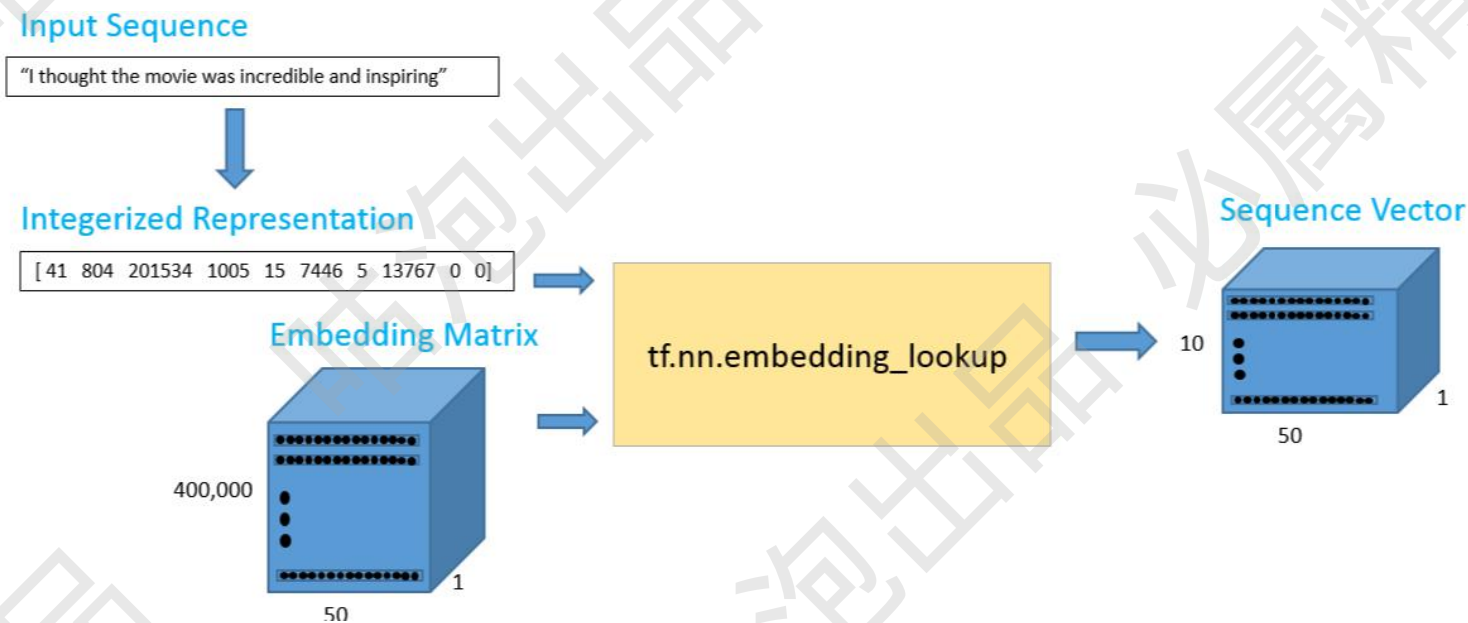


推荐系统

✓ 涉及技术点分析

✎ Embedding方向：如何更好的表示数据，肯定不用one-hot

✎ 隐向量的方法在推荐中几乎无处不用，例如常见的FM及其DeepFM算法



✓ 涉及技术点分析

📌 文本处理方法比较多, LDA,词向量,矩阵分解等套路都能用得上

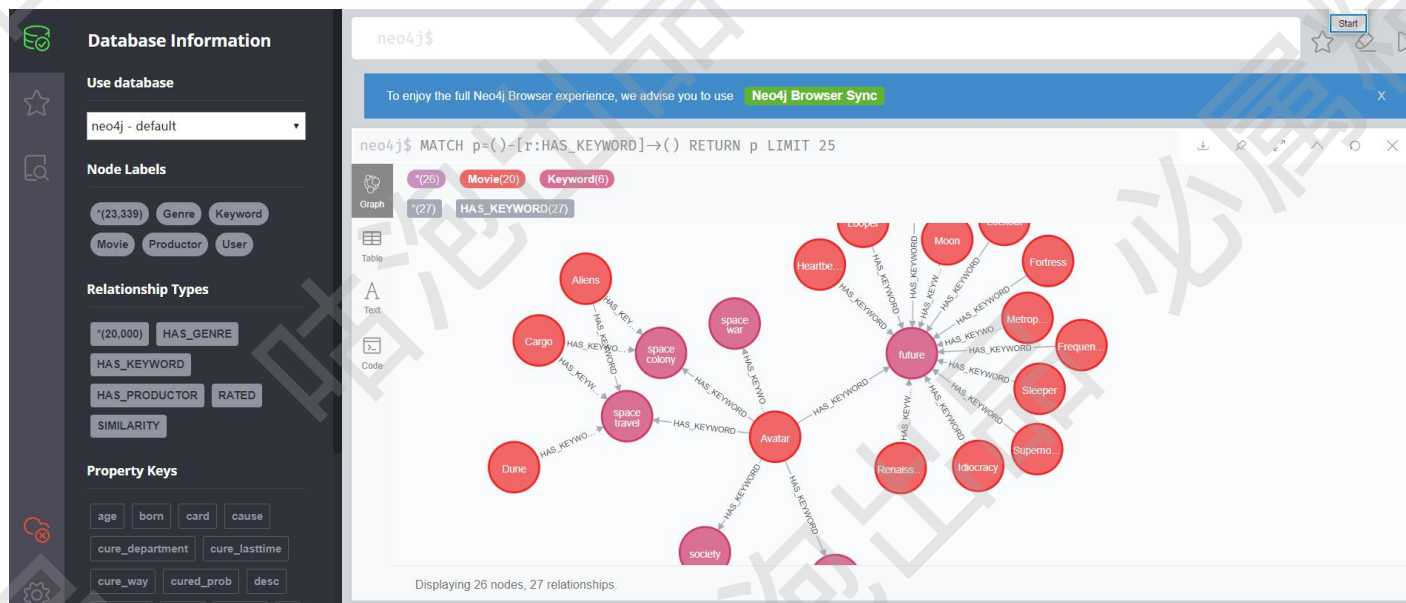


推荐系统

✓ 涉及技术点分析

✎ 知识图谱方向：现在这么火的技术点，推荐中肯定也用到了

✎ 数据越多，越能体现出知识图谱的强大

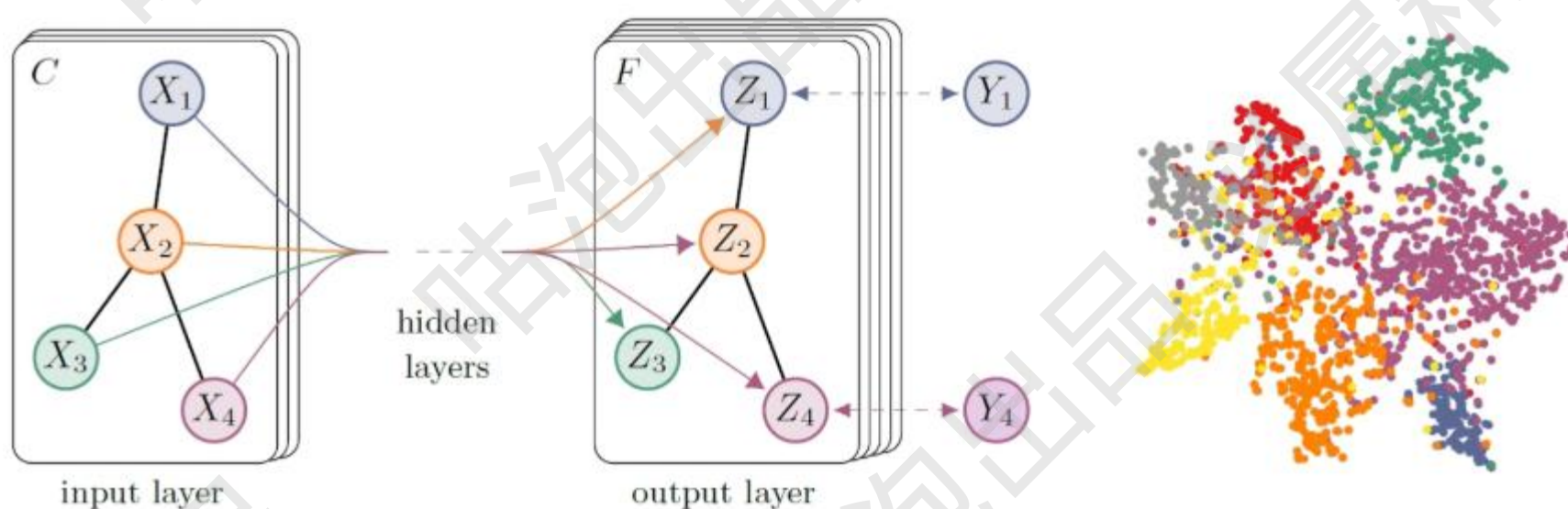


推荐系统

✓ 涉及技术点分析

✎ CV方向：卷积与图卷积，图像与视频数据也是用户行为中的体现

✎ 卷积不仅仅能应用在图像/视频数据中，矩阵数据都可以尝试

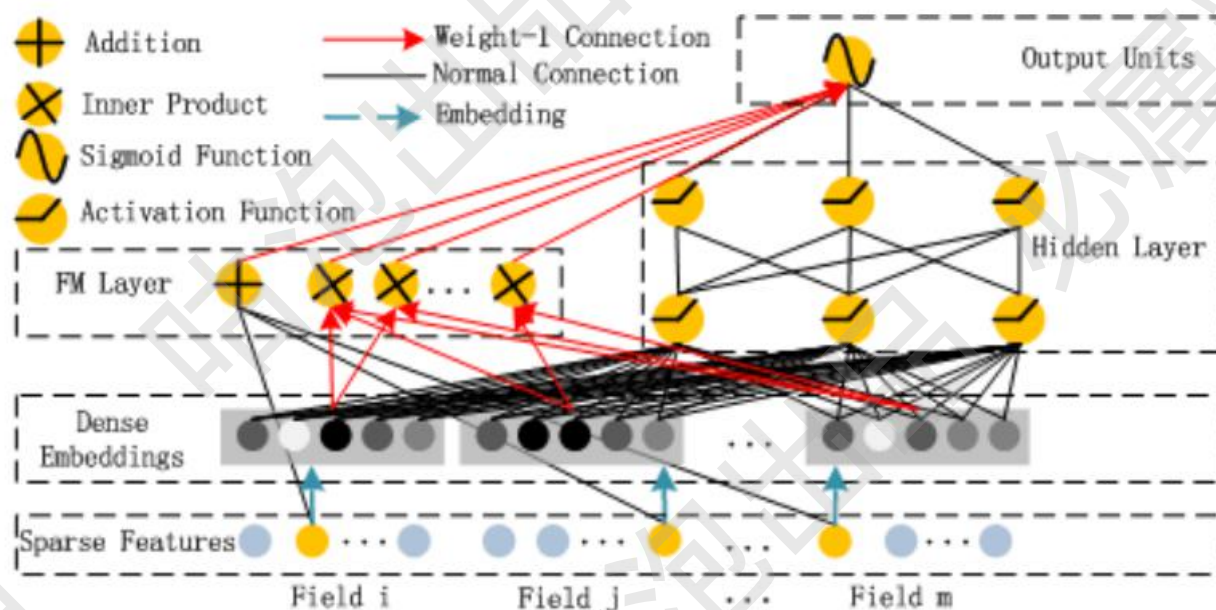


推荐系统

✓ 涉及技术点分析

✎ 特征工程与深度学习方向：如何更好的利用这么多信息？深度学习天生优胜！

✎ 数据维度大，稀疏度高一直都是一个大难题，与深度学习结合能更简单



推荐系统

✓ 为什么需要深度学习

✎ 宏观：传统人工特征工程难度较大，深度学习可以把特征做的更好

✎ 本质：end2end的架构让模型训练起来更容易，项目做起来更简单！

✎ 深度学习更适合NLP与图像数据，符合当下用户行为数据

✎ 但凡看到深度学习，第一感觉应该是这件事做起来没那么麻烦了