



QCon 全球软件开发大会
INTERNATIONAL SOFTWARE
DEVELOPMENT CONFERENCE

BEIJING 2018

《深度学习在视频搜索中的应用》

演讲者 / 刘尚堃

目录

- 1、 视频搜索的挑战
- 2、 深度学习在视频内容理解中的应用——召回
- 3、 深度学习在语义搜索中的应用——语义表征
- 4、 深度学习在排序中的应用——个性化表征

视频搜索的挑战

- 1、非结构化/无组织——召回难度
- 2、短文本/信息不充分——语义难度
- 3、海量短视频——用户选择困难

优酷视频搜索深度学习应用领域：

- 1、基于视频内容理解的召回
- 2、语义模型/语义表征
- 3、个性化表征

内容理解——基于视频内容的召回

基于视频内容召回



基于标题以出现召回不足



基于视频图像以下信息召回

- 1、类目标签
- 2、事件/场景标签
- 3、物体/人物标签
- 4、字幕/OCR/ASR形成NLP标签

内容理解——自动分类技术

构建分类标签

- 目的：输入任意视频，通过内容理解的方法对视频进行类目和标签预测
- 方法：采用CNN+LSTM的sequence-to-sequence prediction的方法
- 效果：
 - 基类目平均准确率>85%



内容理解——事件/行为检测

构建事件/场景标签

事件/行为检测技术

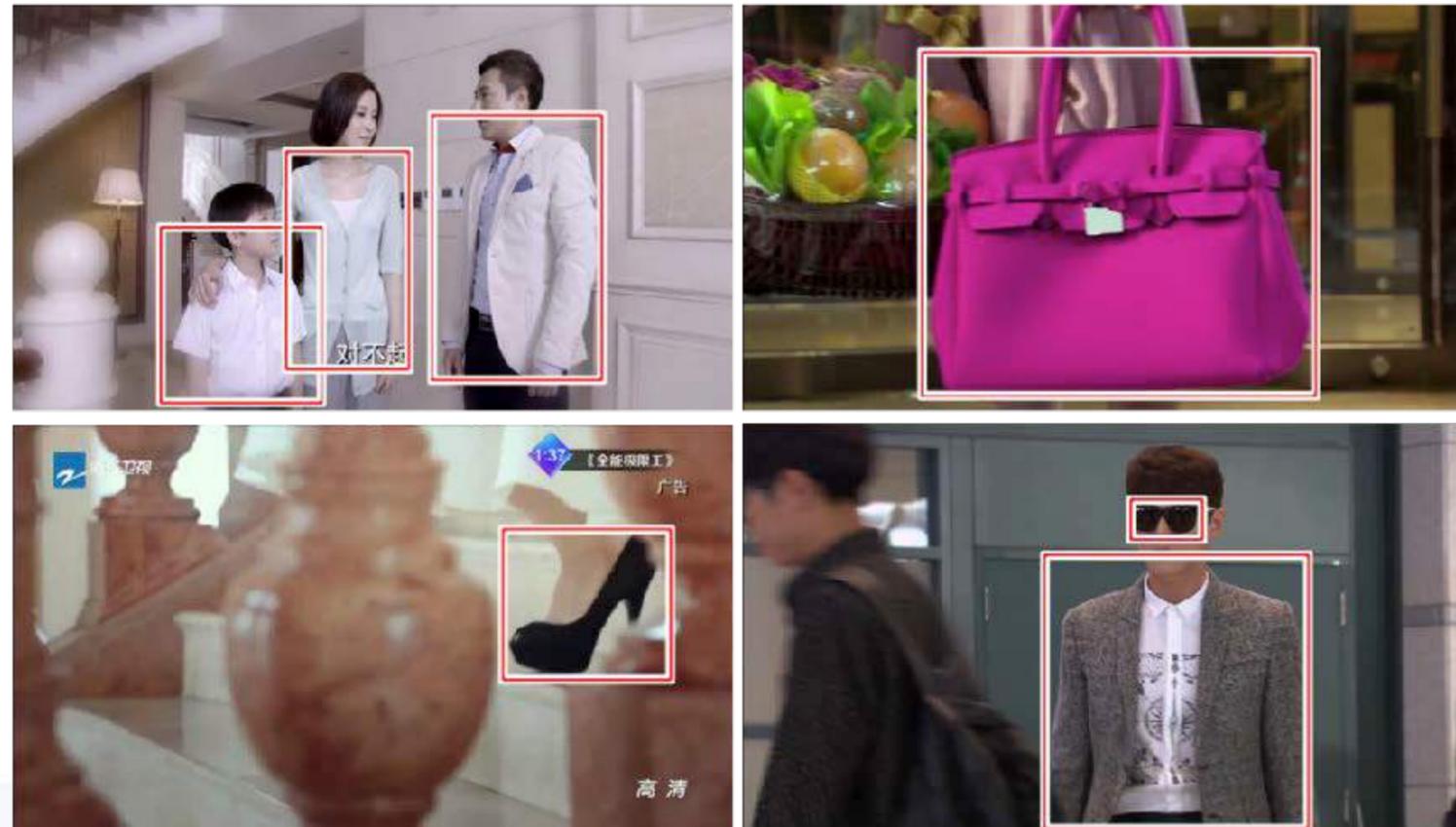
- 目的：给定不定长视频，定位感兴趣行为发生的时间段并给出对应行为类标
- 方法：采取Convolution 3D+Gated Recurrent Units(GRU)算法，结合Single Shot Detector(SSD)框架实现行为检测功能
- 效果：
 - THUMOS 14数据集，mAP=24.1%(IoU=0.5)



内容理解——多目标检测与跟踪

构建物体/人物标签

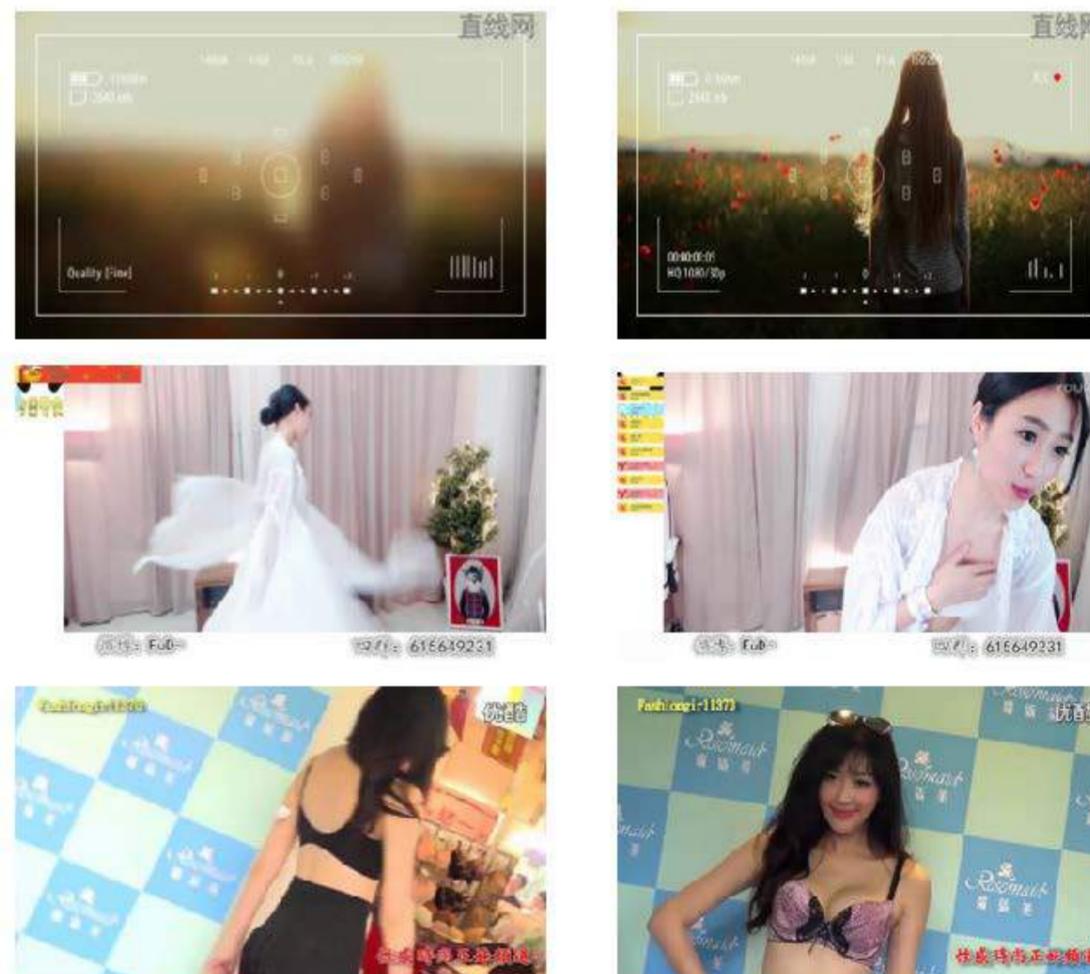
- 目的：定位和识别视频中的特定目标，并在目标生命周期内进行跟踪
- 方法：检测采用Region fully convolution network(R-FCN)的deep learning框架，对于小物体在feature map进行了优化；跟踪采用DCF框架，结合颜色模型，并使用BACF进行候选区域扩充
- 效果：
 - 检测算法在Hollywood Head数据集，mAP=80.41%，高过the state-of-art 7.7%；
 - Brainwash数据集，mAP=88.47%，高过the state-of-art 10.5%。
 - 跟踪算法在TB-100评测集，在实时速度下，precision和success rate效果最好



内容理解——视频智能封面图

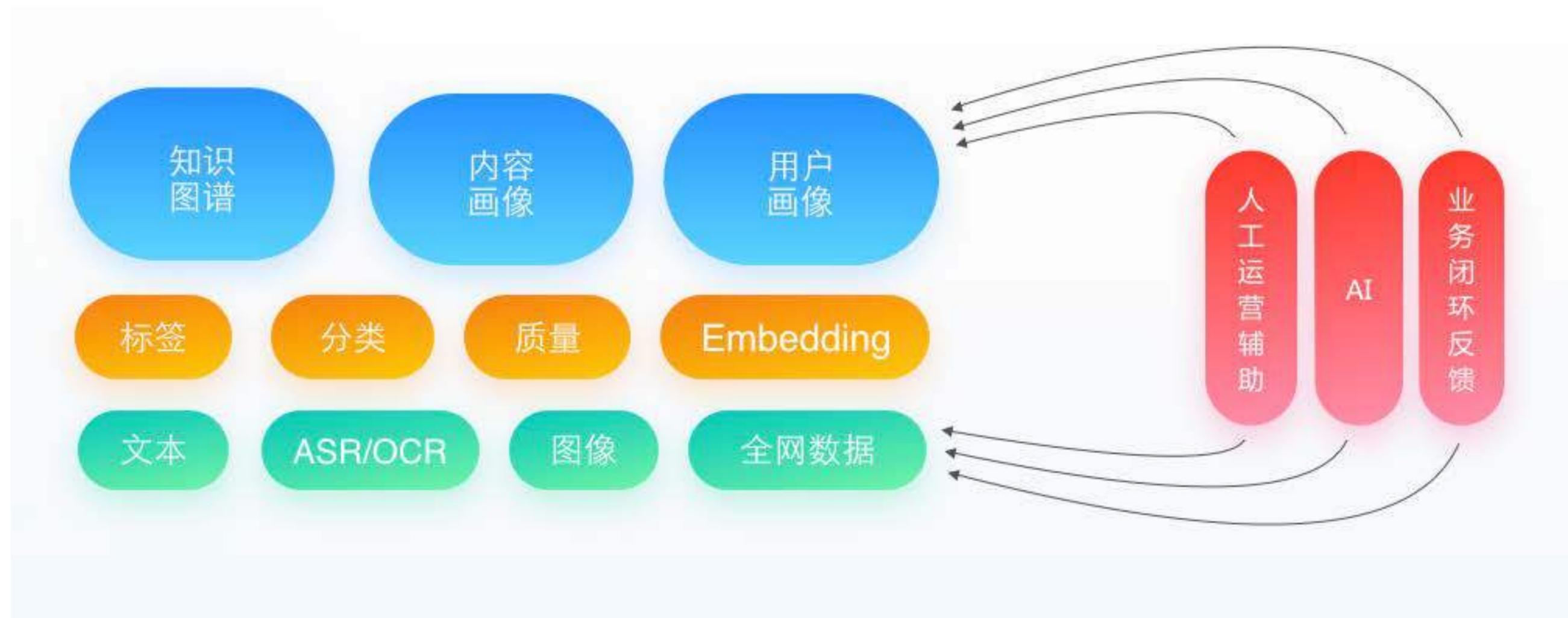
UGC视频智能缩略图

- 目的：通过对视频进行结构化分析，对关键帧、视频镜头进行筛选和排序，选择最优的关键帧、关键片段来作为视频的展示
- 方法：视频智能缩略图采用关键帧提取+MMR优化+美学评分等方法，选择视频中最优关键帧作为该视频的首图。
- 效果：
 - 人工评测优酷原始与封面图算法（仅评测可对比出优劣之分的数据）：算法优，占比72.2%



UGC视频智能缩略图示例
左边：优酷线上原图 右边：算法生成

内容理解——总结



语义模型

- 测试集：语义人工标注ground truth

- 目前最高NDCG : 0.9x

- 固定数据尝试不同模型：

- 双向LSTM+Attention 0.9x

- BiGRU dropout 0.8x

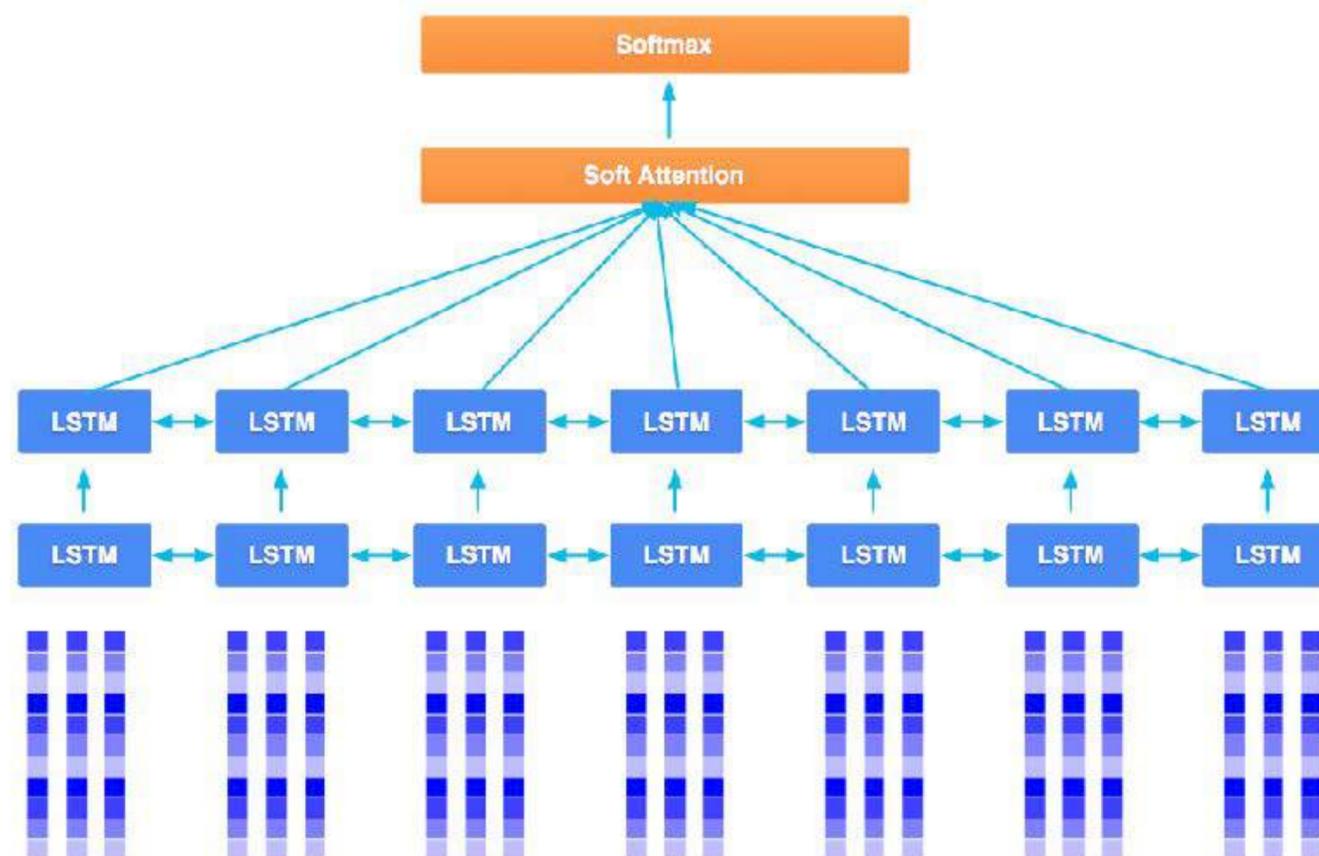
- 固定模型尝试term embedding初始化方式：

模型	初始化方式	长尾Query NDCG
bilstm+attention	随机	基线
bilstm+attention	神马模型输出embedding	低于基线1%
bilstm+attention	UC、豆瓣Fasttext embedding	高于基线1%

- 标注集：针对多次人工check标注QU（训练集测试集比例7：3）

加入语义特征后：NDCG 提升1%绝对值。

- 特征重要度仅次于匹配度特征。



语义模型——初始化方式fasttext

original_id	new_version	old_version
黑科技	VR	Vive
黑科技	虚拟现实	超级
黑科技	消费级	虚拟现实
黑科技	酷炫	炫酷
黑科技	EUI	游戏主机
黑科技	VR游戏	Game b r y o
黑科技	游戏主机	A V A T E R
黑科技	微型投影机	消费级
黑科技	科技产品	神器
黑科技	可穿戴设备	酷炫
黑科技	冰炭不投	蝴蝶机
黑科技	炫酷	高通骁龙

original_id	new_version	old_version
宝宝	小宝宝	小宝宝
宝宝	宝贝	婴儿
宝宝	婴儿	宝贝
宝宝	爸爸妈妈	喂奶
宝宝	妈咪	宝宝洗澡
宝宝	小宝贝	认生
宝宝	宝宝洗澡	王尚江
宝宝	喂奶	宝宝健康
宝宝	妈妈	爸爸妈妈
宝宝	小孩子	妈咪
宝宝	拍嗝	母乳喂养
宝宝	月龄	月龄
宝宝	淘气	拍嗝
宝宝	宝宝健康	哭闹
宝宝	宝贝儿	安抚奶嘴

original_id	new_version	old_version
芳华	匆匆那年	花事
芳华	杜丽娘	风华
芳华	励志剧	流光
芳华	情深深雨蒙蒙	回眸
芳华	英雄本色	流年
芳华	周党镇	暗香
芳华	花谢花飞花...	春色
芳华	红颜	年华
芳华	青丘狐传说	韶华
芳华	金陵十三钗	干贝汤
芳华	女人花	落花
芳华	春风十里不...	花影
芳华	冰雪公主	洗尽铅华

语义模型——举例

query	vdo_title	label	语义预测	非语义预测	模型解释
2017香港小姐竞选	TVB《2017国际中华小姐》竞选佳丽学仪态走猫步无时无刻加紧练习	2	2		1 embedding初始化
big笑工坊唐唐脱口秀	【牛人】Big笑工坊 第一季：唐唐神吐槽:最作死的女神 184	3	3		2 语义
中国达人秀震惊全场	中国好声音李安19岁参加澳洲达人秀时震惊全场的表演	2	2		3 长短距离依赖
潜规则女秘书	职场大尺度虐恋激情电影《错爱，爱错》（性感女秘书欲望潜规则男上司）	1	1		2 长短距离依赖
日本广岛原子弹爆炸	100810_廣島_原子彈爆炸後的日子_寰宇地理_CETV	3	3		2

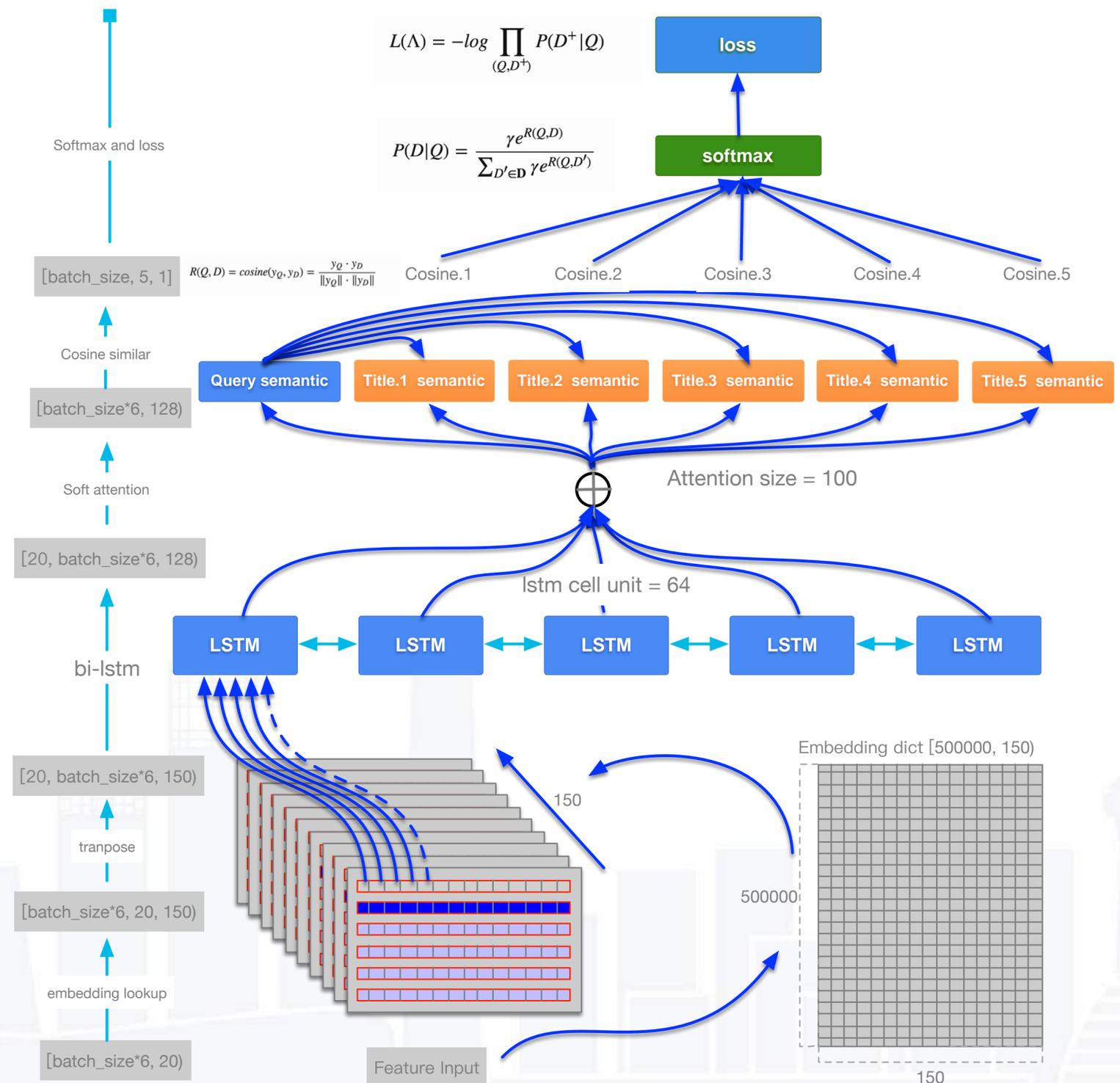
语义模型——总结

用户体验优化

- 在长尾query和语义层面实现了特征增益, 在长尾query相关性上有较大改善
- ground truth测试集NDCG提升1%

技术创新突破

- 基于FastText Vector无监督embedding
- 细粒度+字切分 优酷title语料覆盖度99.0%
- 训练数据量 billion级别 模型参数量千万级别
- bi-LSTM + attention
- 基于pai-tensorflow的分布式训练



语义模型——总结

The image displays four screenshots of a video search interface, likely from a mobile application. Each screenshot shows a search bar at the top with a search term, followed by a list of video results. The results are organized into sections like '精选' (Selected) and '自频道' (My Channel). The first screenshot shows search results for '适合宝宝看的电影:mtest:8'. The second screenshot shows search results for '适合宝宝看的电影:mtest:2'. The third screenshot shows search results for '怎么滑雪比较酷:mtest:8'. The fourth screenshot shows search results for '怎么滑雪比较酷:mtest:2'. Two specific video results are highlighted with red boxes: '最适合儿童看的八部电影' in the first screenshot and '爽!滑雪真酷' in the third screenshot. A red 'X' is also visible over a video result in the fourth screenshot.

Screenshot 1: 适合宝宝看的电影:mtest:8

- 最适合儿童看的八部电影
上传者: 大宇知识讲堂 2月前
- 10部适合孩子学英语的动画电影,寒假陪孩子一起看
上传者: 灵灵大杂烩 26天前
- 宝宝最爱看的亲子视频 适合孩子看的动画片 儿童启蒙动画视频
上传者: 超牛视频推荐啊 2年前
- 看看老外是怎么切苹果的, 超级有爱心, 适合给宝宝做
上传者: 芒果之文 2月前
- 陪宝宝看电影
上传者: 凤帆TV 12天前
- (爆笑系)海绵宝宝和派大星看电影
上传者: 小迪 3天前
- 9个月宝宝看电影
上传者: fanchengming321 1月前

Screenshot 2: 适合宝宝看的电影:mtest:2

- 4分钟看完恐怖电影《鬼娃回魂4》,鬼娃楚奇和鬼娃新娘生出宝宝
上传者: 丁丁解说 2天前
- 海绵宝宝和派大星看电影,坏人聚会要干什么
上传者: 热血动漫 24天前
- 海绵宝宝和派大星看恐怖电影不敢睁开眼睛,就看了几遍!哈哈
上传者: 海绵动画得瑟哥 1月前
- 最适合儿童看的八部电影
上传者: 大宇知识讲堂 2月前
- 适合胆小者看的恐怖电影解说,9分钟带你看完烧脑恐怖电影《骇故事之女舍惊魂》
上传者: 小海讲电影 7月前
- 萌萌哒的电影 适合妈妈带着宝宝一起看
上传者: 福乐 8年前
- 赶紧收藏!10部适合中小學生看的电影(附最全观影攻略)
上传者: 椰椰课堂 20天前

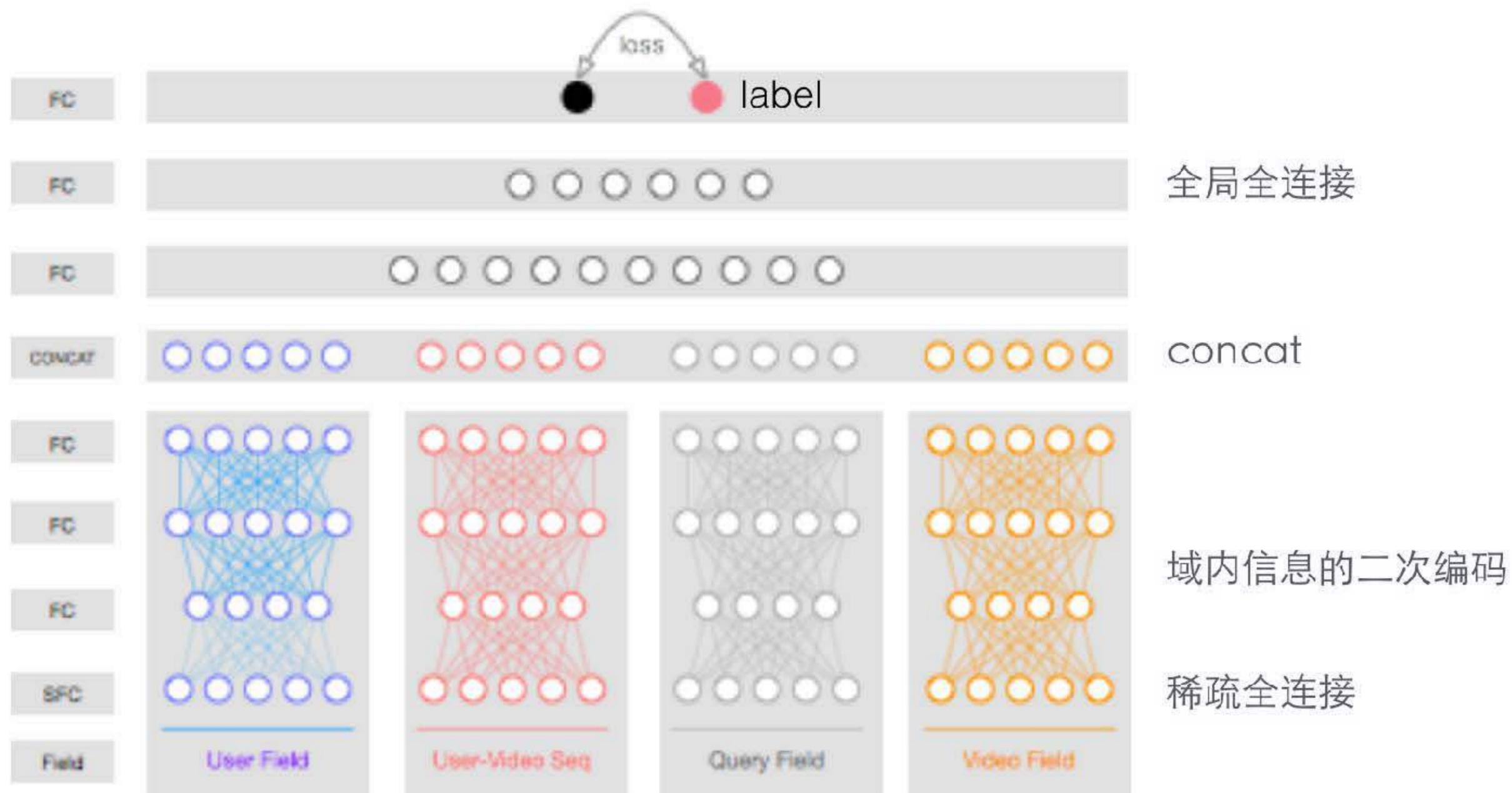
Screenshot 3: 怎么滑雪比较酷:mtest:8

- 单板滑雪太酷了!超酷!
上传者: 前端体育 1月前
- 我的世界怪物学院《高台滑雪》,苦力怕的咸鱼单板玩的真酷
上传者: 方块菜鸟 19天前
- 双板滑雪PK单板滑雪谁更酷
上传者: 小马撒欢爆笑短视频 3月前
- 爸爸和女儿滑雪板在一起,和他们怎么做会让你大吃一惊!
上传者: sleiman 1年前
- 顽酷滑雪工厂-小会员生日会
上传者: 体育盛君 19小时前
- 吴亦凡化身滑雪酷boy 王思聪评论:去戴个头盔
上传者: 天天汇娱乐 17天前
- 爽!滑雪真酷
上传者: 等一个晴天15532441 4年前

Screenshot 4: 怎么滑雪比较酷:mtest:2

- 单板滑雪太酷了!超酷!
上传者: 前端体育 1月前
- 我的世界怪物学院《高台滑雪》,苦力怕的咸鱼单板玩的真酷
上传者: 方块菜鸟 19天前
- 双板滑雪PK单板滑雪谁更酷
上传者: 小马撒欢爆笑短视频 3月前
- 爸爸和女儿滑雪板在一起,和他们怎么做会让你大吃一惊!
上传者: sleiman 1年前
- 穷玩车富玩表,2017有哪些比较酷的智能手表?
上传者: 康智 2月前
- 顽酷滑雪工厂-魔毯无限雪道单板练习
上传者: 体育盛君 17天前
- 吴亦凡化身滑雪酷boy 王思聪评论:去戴个头盔
上传者: 天天汇娱乐 17天前

排序应用——个性化表征



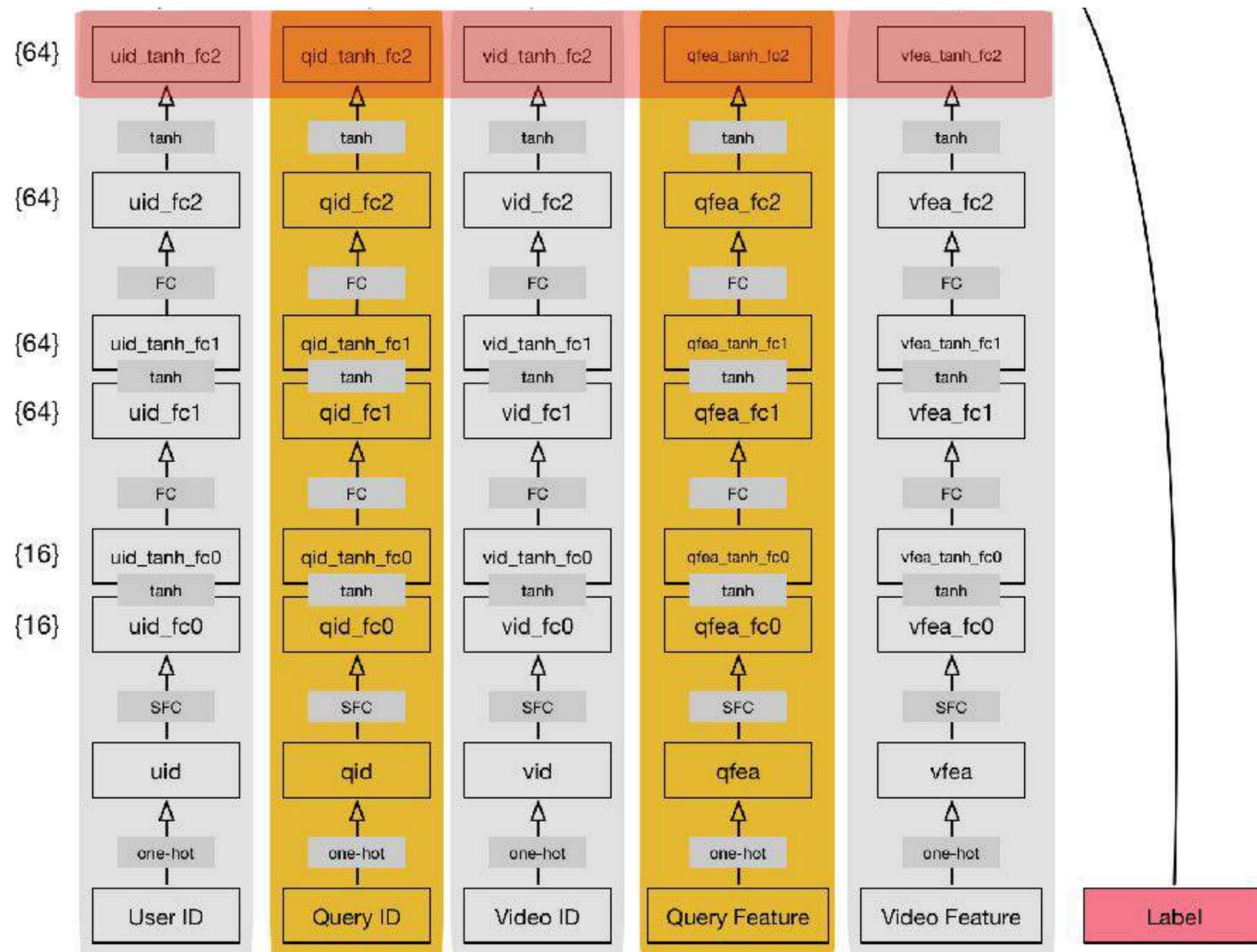
排序应用——个性化表征

1、特征域划分及编码

- query user video id域 统计域 用户观看序列 标签兴趣 文本
- 超高维的稀疏编码来表征独立个体。上亿维。深度特征的组合表达能力
- 按特征的重要度和关联性分域

2、特征降维 压缩

- 在线计算挑战：特征维度高 模型存储空间大，离线训练计算时间成本高，在线实现资源占用高，前向网络计算不能满足RT要求
- 特征分域
- 随机编码
- 挂靠编码
- 抽样技术



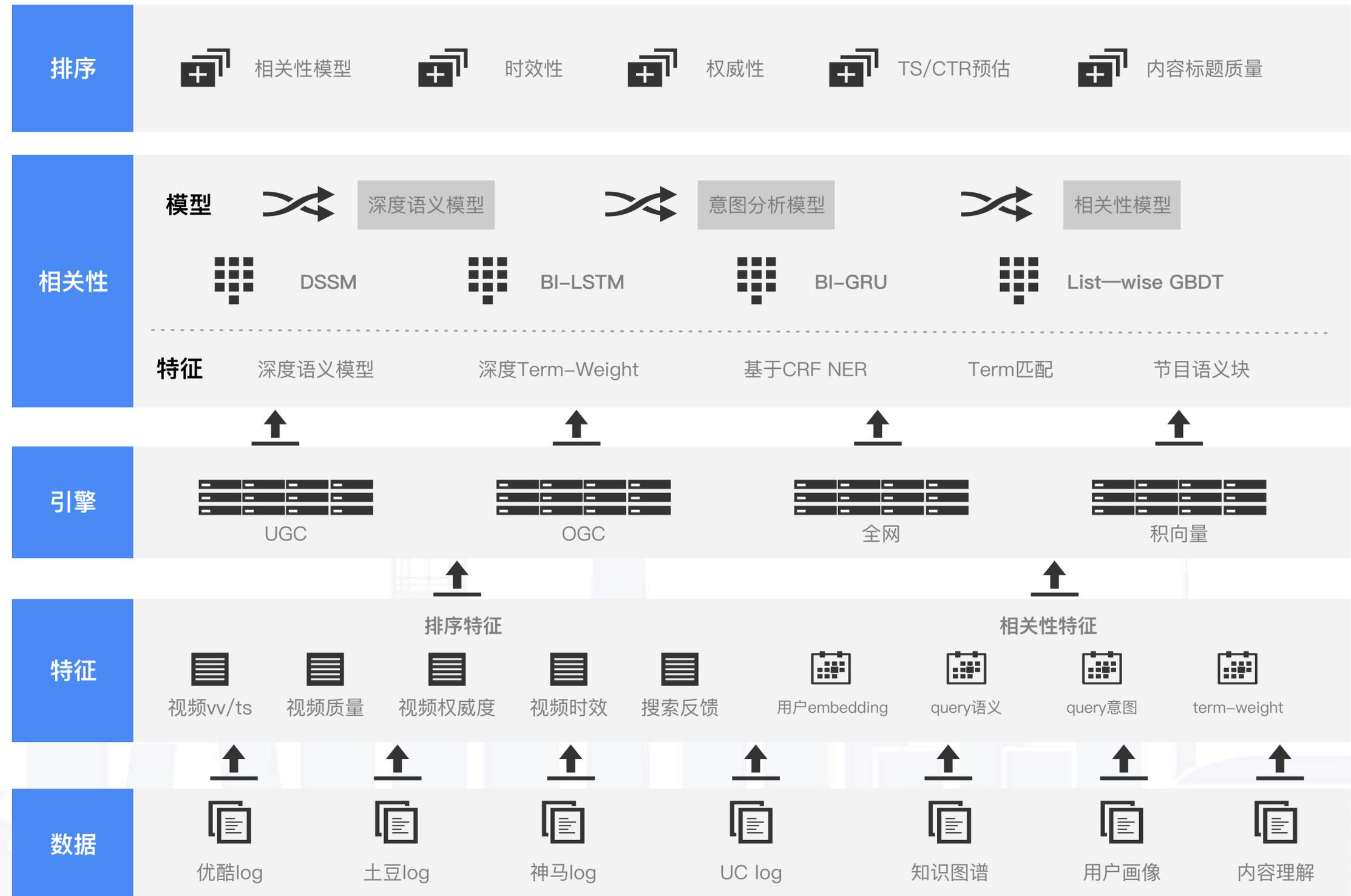
总结

基于高效内容分发的场景应用

- 基于深度语义的相关性应用
- 基于表征学习的TS预估
- 新一代积向量召回引擎

超大规模，高维稀疏训练样本

- 深度语义：训练样本X billion*X万级（稀疏）
模型参数亿+ 迭代次数X million步
- 表征学习：训练样本X billion*X million（稀疏）
模型参数 X亿+ 迭代次数 Xmillion





关注QCon微信公众号，
获得更多干货！

Thanks!



主办方 **Geekbang** & **InfoQ**
极客邦科技