

AI平台化：助力提升大电商产品力

演讲人：于磊

过去，OTA/电商平台产品力：

过去：产品力

1. 产品数量多
2. 追求绝对低价

结果

不健康的平台生态

1. 平台：只顾着跑马圈地
2. 资方、平台和供应商：囚徒困境
3. 用户：得不到有品质的产品和服务；无法消费升级

大OTA/电商平台新产品力定位：

过去：低价v.s融资

1. 产品数量多
2. 追求绝对低价

重点

现在：良性平台生态->品质 v.s 成本

1. 有品质/吸引力的产品
2. 有吸引力的产品信息
3. 有竞争力的产品服务

回归初心…

旅行

只为对**未知**世界的好奇

OTA产品力：产品和产品信息展示，对**不熟悉**它的客人，有足够**吸引力**

AI探知：有品质/吸引力的产品

AI运营：有吸引力的产品信息

客服AI：有竞争力的产品服务

+
基础AI能力的引入机制



灵石. 崇宁堡温泉酒店

北方的深宅大院



澳大利亚·海豚岛天阁露玛度假村
酒店海滩，喂野生海豚



阳朔1Q84·宿

An aerial photograph of a modern hotel terrace built on a rocky cliff overlooking the ocean. The terrace features a long bar with a blue backlit counter, several lounge areas with yellow and grey seating, and glass railings. Waves are crashing against the rocks below the terrace. The sky is clear and blue.

巴厘岛金巴兰 RIMBA 酒店



巴厘岛·马拉河野外巴厘旅馆

楼下和窗外就是野生动物



澳门·新濠影汇酒店 (Studio City Hotel)
8字摩天轮酒店



丽江·十里雪山360度高尔夫野奢庄园 酒店

AI探知：有吸引力/品质的产品

Ctrip 整站搜索词

Ctrip 整站评论

Ctrip 用户QA

旅游类公众号

各类微博数据

互联网数据

...

AI理解

多维度统计

当月网红民宿

综合排名	民宿名称	当月指标	前期指标	综合指数
1	大理蓝雀艺术精品客栈	19		1
2	苏州太湖尘外小院客栈			7
3	苏州太湖西山岛香樟小院精品民宿		3	
4	深圳西涌故事客栈		2	
5	莫干山花木深度度假民宿			
6	武夷山福泰源民宿			
7	杭州慕仁露营·太阳谷·拾伍间民宿			
8	广州宛若故里·稻田里民宿			3
9	阳朔青雲閣民宿			2
10	雅安碧峰峡天然居主题客栈	2		2
11	乡伴余姚树蛙部落民宿	4		2
12	时光码头江景民宿(重庆解放碑洪崖洞店)	2		2
13	丽江龙景轩精品度假客栈	3	1	2
14	凤凰无梦高端度假民宿	2	5	2
15	苏州太湖小园度假客栈	7	1	2
16	厦门三个厦大人kiwi民宿			2
17	乌镇半望山·问心民宿	2		2
18	聆海沐月客栈	2		19
19	上海愿丽精品民宿	1		19
20	大理双廊村野海晏度假客栈	1	25	19

AI探知：当地旅游活动/节日

- Ctrip 整站搜索词
- Ctrip 整站评论
- Ctrip 用户QA
- 旅游类公众号
- 各类微博数据
- 互联网数据
- ...



AI探知：有品质/吸引力的产品

AI运营：有吸引力的产品信息

客服AI：有竞争力的产品服务

+
基础AI能力的引入机制

AI平台：挖掘有吸引力的产品信息

普洱小熊猫庄园： AI探知的卖点(喂小熊猫、猴子)

1.00000 000	<div>喂食</div> <div>小熊猫</div>	<p>因入住体验非常棒，环境很好，去哪里都可以让前台安排电瓶车送，工作人员态度都很好，前台就可以看到小熊猫，每天也有活动是酒店专门安排的可以喂食小熊猫和犀牛///</p>
1.00000 000	<div>喂</div> <div>小熊猫</div>	<p>坪森林公园里面，环境特别好，有很多小动物小蘑菇，鸟叫蝉鸣不绝于耳，虽然酒店床品等硬件不够五星，但是我觉得它的服务以及环境真的太加分了，森林体验师小崔带我们喂了鹿喂了小熊猫猴子等动物，小朋友们玩得特别开心，一段非常难忘的旅程，给了我们一段非常难忘的旅程，特别感谢小崔的耐心讲解///</p>



庐境西溪酒店： AI探知的卖点(网红玻璃船、网红泳池)

19517372	杭州庐境西溪酒店	1.8764236121 835907	<div>网红游泳池</div> <div>网红玻璃船</div>
----------	----------	------------------------	-----------------------------------



AI平台：整合营销/IP整合活动

Ctrip 大数据

Ctrip全站数据
互联网全网数据

携程AI：10月营销热点-枫叶
10月IP热点-悼念金庸



AI自动化跨域流程

枫叶景点

金庸IP景点

酒店产品

团队游产品

自由行产品

门票产品

AI跨域发现整合的产品

团队游产品：悼念金庸的IP活动



AI精准化投放

金庸粉
关注赏枫的公众
有意去产品目的地的用户

供应商更新信息/生成活动页面



酒店产品：由枫叶景点的跨域营销



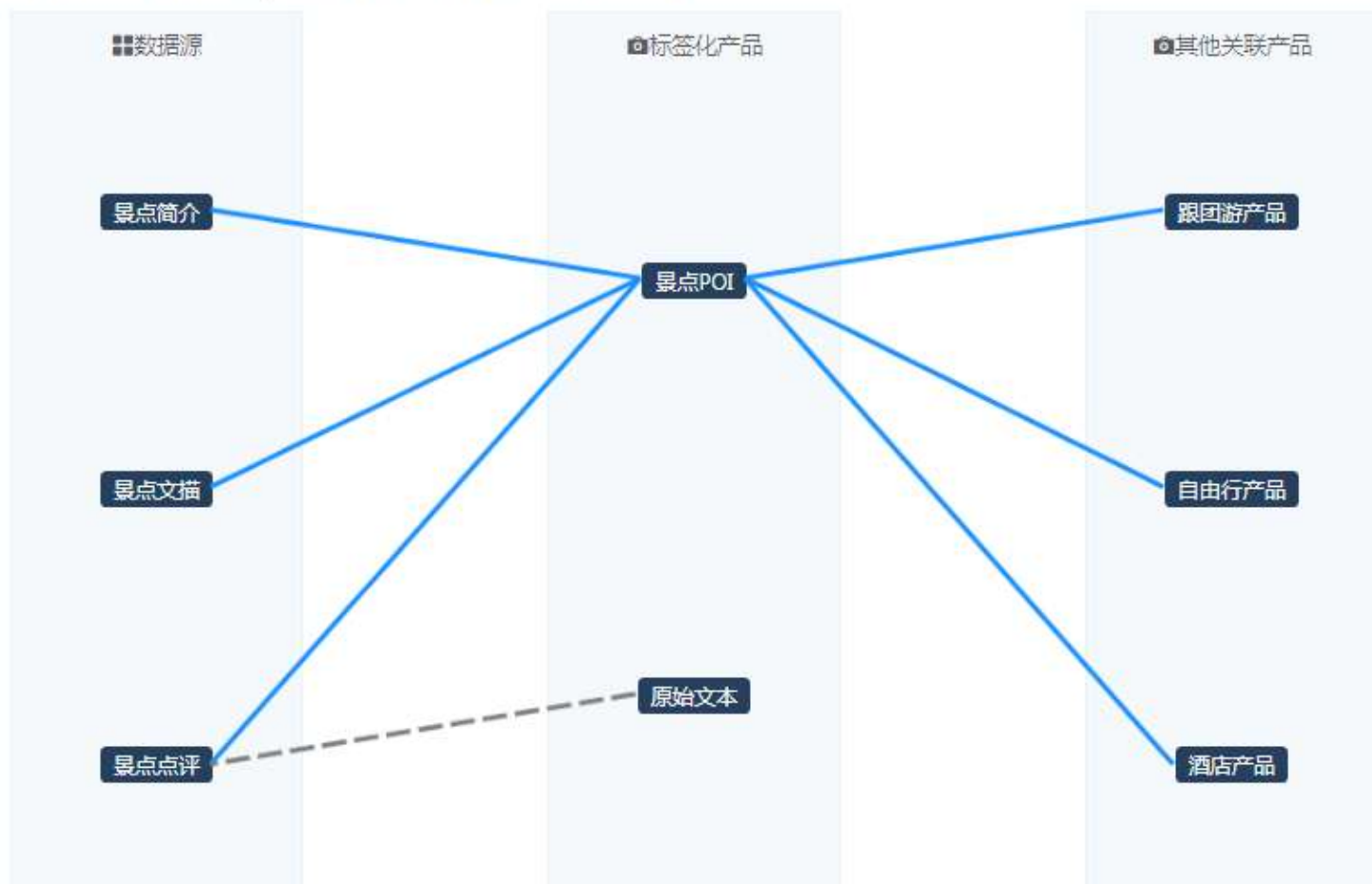
AI平台化：跨域的自动化整合流程

AI挖掘产品预览

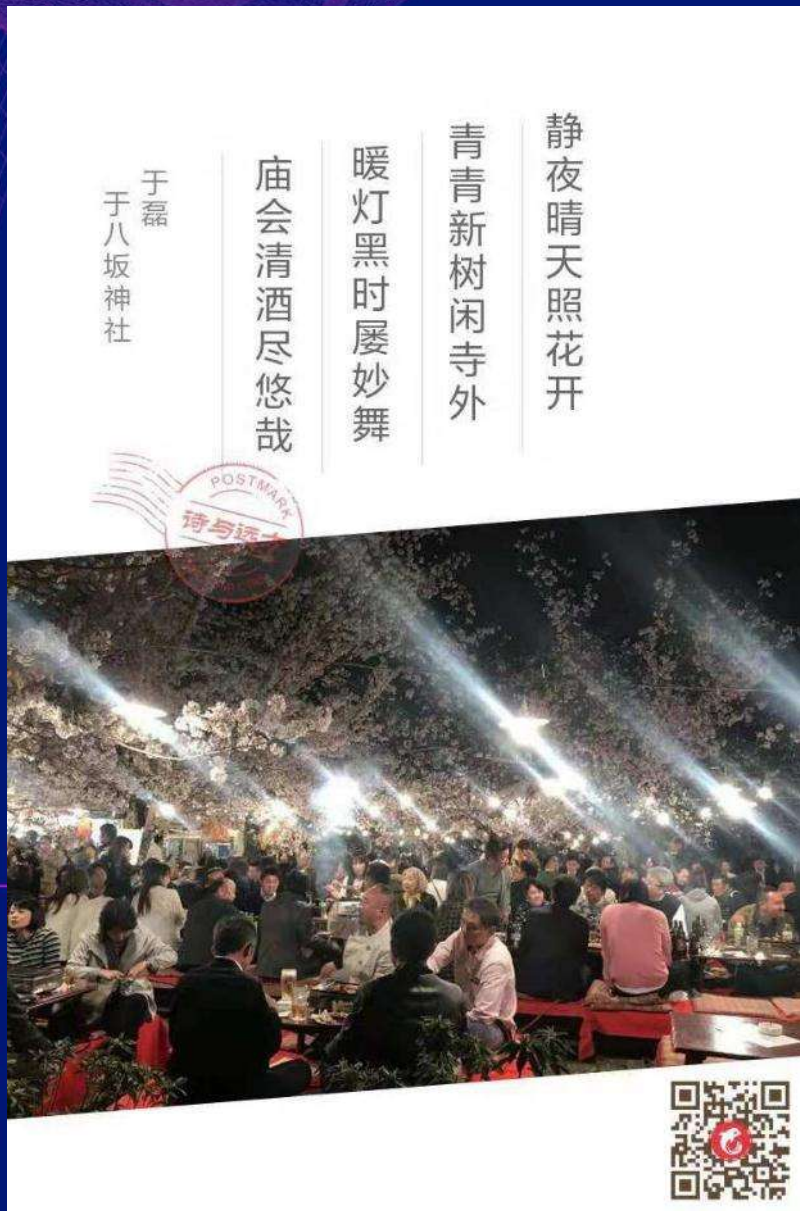
Score	产品	网址
	天龙八部影视城	
	洗马潭大索道	
	崇圣寺三塔文化旅游区	
	象山影视城	
	苍山	
	桃花岛风景区	
	大理古城	
	光明顶	
	千文幽谷	
	襄阳古城	

标签 枫叶 金庸系列IP

请选择最终生效的数据源、需要绑定的标签化产品以及其他关联产品



AI平台化：产品文字内容AI生成



hotelid: 22987483
 标签:
 洲际;洲际酒店_1_酒店_住_机构名称
 鲤城;刺桐城;温陵;泉州_0.1_城市名_现代地理名称_地理名称_泉州_PoemKnowledgeBaseCity
 丰海路_0.1_街道_现代地理名_地理名称_丰海路_PoemKnowledgeBaseStreet
 高端_1_酒店消费等级_酒店综合评价_酒店属性
 强烈推荐_1_酒店用户评价_酒店综合评价_酒店属性
 公园_1.0_林_树_植物_泉州森林公园_NearbyAttractions
 海印寺_1.0_寺_宗教建筑;建筑群_建筑_海印寺_NearbyAttractions
 推荐理由:

在洲际，坐拥绵延公园，享受绿色静谧的度假时光

hotelid: 22824733
 标签:
 十里桃花;十里桃花旅社_1_酒店_住_机构名称
 皇城;帝都;京城;北京;燕京;京薊;北平;首都;薊城_0.1_城市名_现代地理名称_地理名称_北京_PoemKnowledgeBaseCity
 北甲地路_0.1_街道_现代地理名_地理名称_北甲地路_PoemKnowledgeBaseStreet
 经济_1_酒店消费等级_酒店综合评价_酒店属性
 强烈推荐_1_酒店用户评价_酒店综合评价_酒店属性
 博物馆_1.0_博物馆_NULL_建筑_辽金城垣博物馆_NearbyAttractions
 儿童_1.0_乐园_大建筑_建筑_保利儿童乐园_NearbyAttractions
 游乐园_1.0_乐园_大建筑_建筑_峡谷达人儿童游乐园_NearbyAttractions
 拓展_1.0_林_树_植物_乐仕堡儿童运动拓展中心（草桥店）_NearbyAttractions
 推荐理由:

皇城的十里桃花旅社，经济舒适，坐拥繁华的都市风景

个性化生成：产品文本内容

AI平台化：对产品图片/文本的流程化整理

400	400::HTL::guide	下雪了，撒欢去！	全球户外滑雪酒店大搜罗	洁白的积雪，五彩的装...		2018-01-10 16:00	945	世界文化遗产-河流	可用	HotelProductInfo	400	1	描写城市/榜单文章	同步	基础数据 特征数据 统计数据
422	422::HTL::guide	去成都耍，肯定要住这儿喽	蓉城，一座来了就不想走的城市	“来来来，眼哥，凤姐...		2018-01-28 16:00	965	潜水-世界文化遗产-动物园	可用	HotelProductInfo	422	1	描写酒店文章	同步	基础数据 特征数据 统计数据

AI对产品图片的识别/盘点

标题

下雪了，撒欢去！

摘要

洁白的积雪，五彩的装备，手持长杆，脚踏滑板，从高山呼啸而下——滑雪，大约是冬季里最具魅力的娱乐项目。

文本内容

<p>人工雪场固然方便，然而，大自然的白色馈赠才是滑雪爱好者们每年都翘首以盼的终极目标，那些围绕着全球著名滑雪圣地的优质酒店，更是在酣畅驰骋后快速回血的首选。</p>

文章首图



▶ 四季属性识别：冬季:7.0

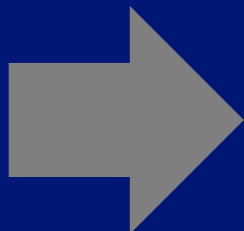
▶ 人群属性识别：不确定:1

▶ 文章首图识别：滑雪:1.0

AI对文字类的识别/盘点

文章标签挖掘：			
tag	tag分数		
雪景	5.8	标题	下雪了///
		摘要	洁白的积雪///
		摘要	从高山呼啸而下——滑雪///
滑雪	3.8	内容	人工雪场固然方便///
			大自然的白色馈赠才是滑雪爱好者们每年都翘首以盼的终
			那些围绕着全球著名滑雪圣地的优质酒店///

AI平台化：产品个性化->产品售卖信息的个性化



原产品售卖/个性化推荐

AI内容化后的售卖/个性化推荐

AI平台化：产品个性化->产品售卖信息的个性化



同一张机票，不同兴趣的人，不同的产品信息

AI平台化：产品个性化->产品售卖信息的个性化

用户画像
(个性化标签)

升级

用户意图

升级

AI兴趣图谱

2015年

T+1天级更新
用户人身上的标签

2016年

T+0实时更新
用户行程和目的地
级别的行为特征

2018年

T+0实时更新
用户分行程分BU的AI特征
目的地多级Category图谱

消费兴趣图谱：

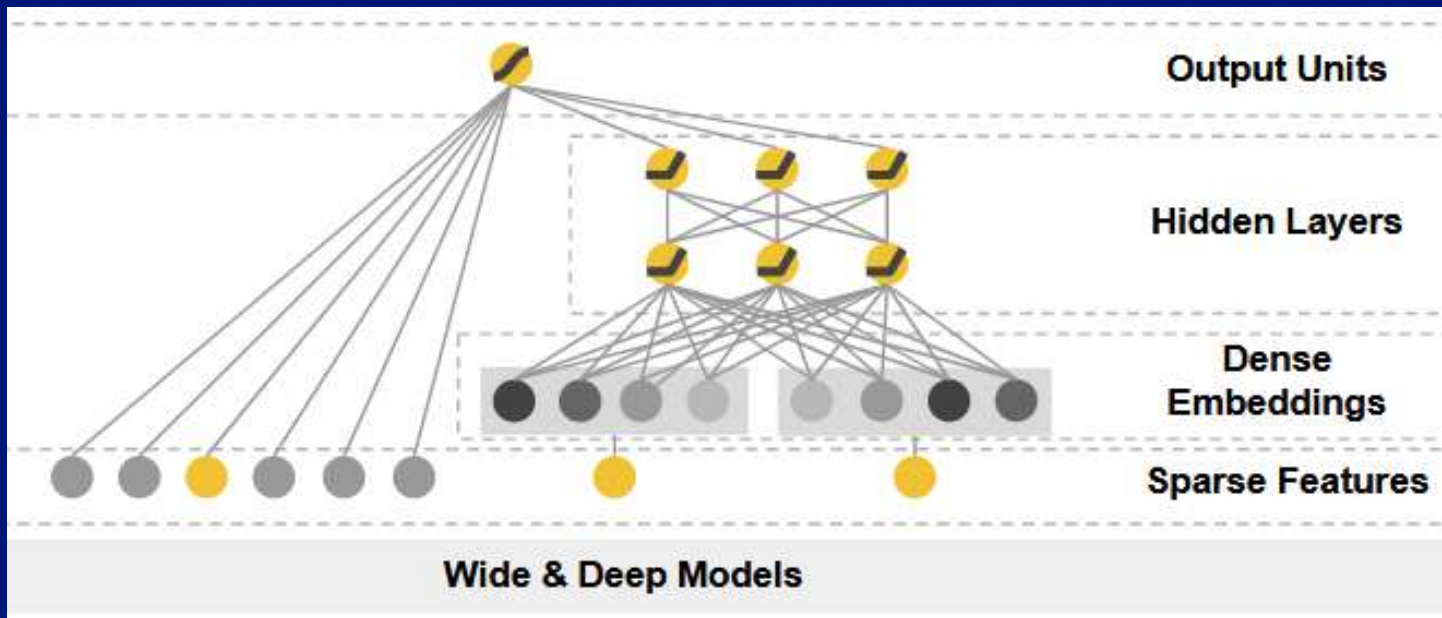
某上海用户在长三角、亚洲、北美的出行人偏好、当前偏好（亲子、父母、情侣、团建、）

某上海用户在伦敦、西安、苏州的住宿偏好（星级、装修、设计师、历史、宠物）

某用户或上海用户群的亲子、爸妈、情侣的Topic偏好、目的地偏好、其它属性偏好

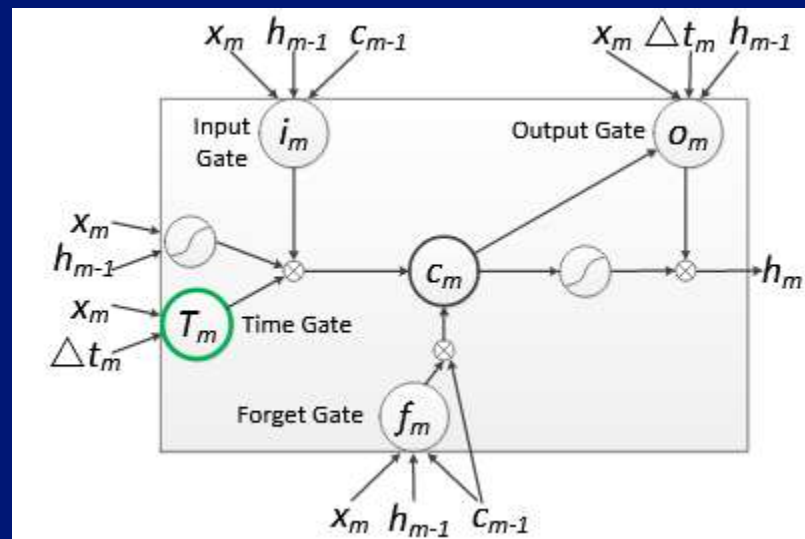
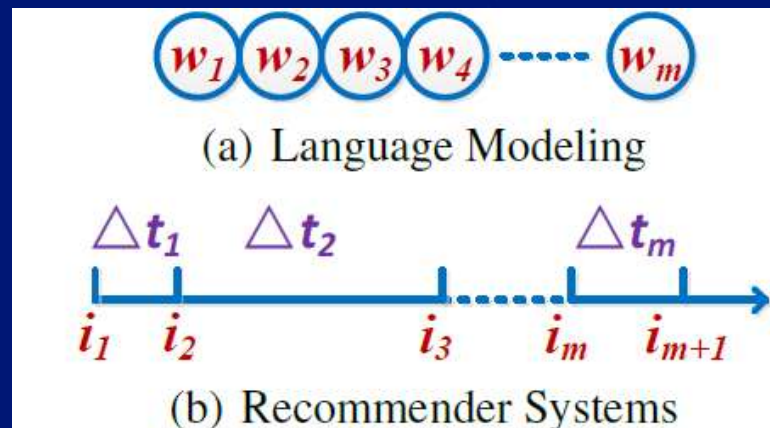
Offline Recommender: Item Ranking via Deep Learning

- Ranking products lists by click through rate(CTR)
- Using Wide & Deep model , Predict CTR value for each user
 - LR + MLP + Embedding
- Features : user preference , browse/order behavior , product properties and statistics ,context information etc.
- Better AUC than LR,FFM on real business data.

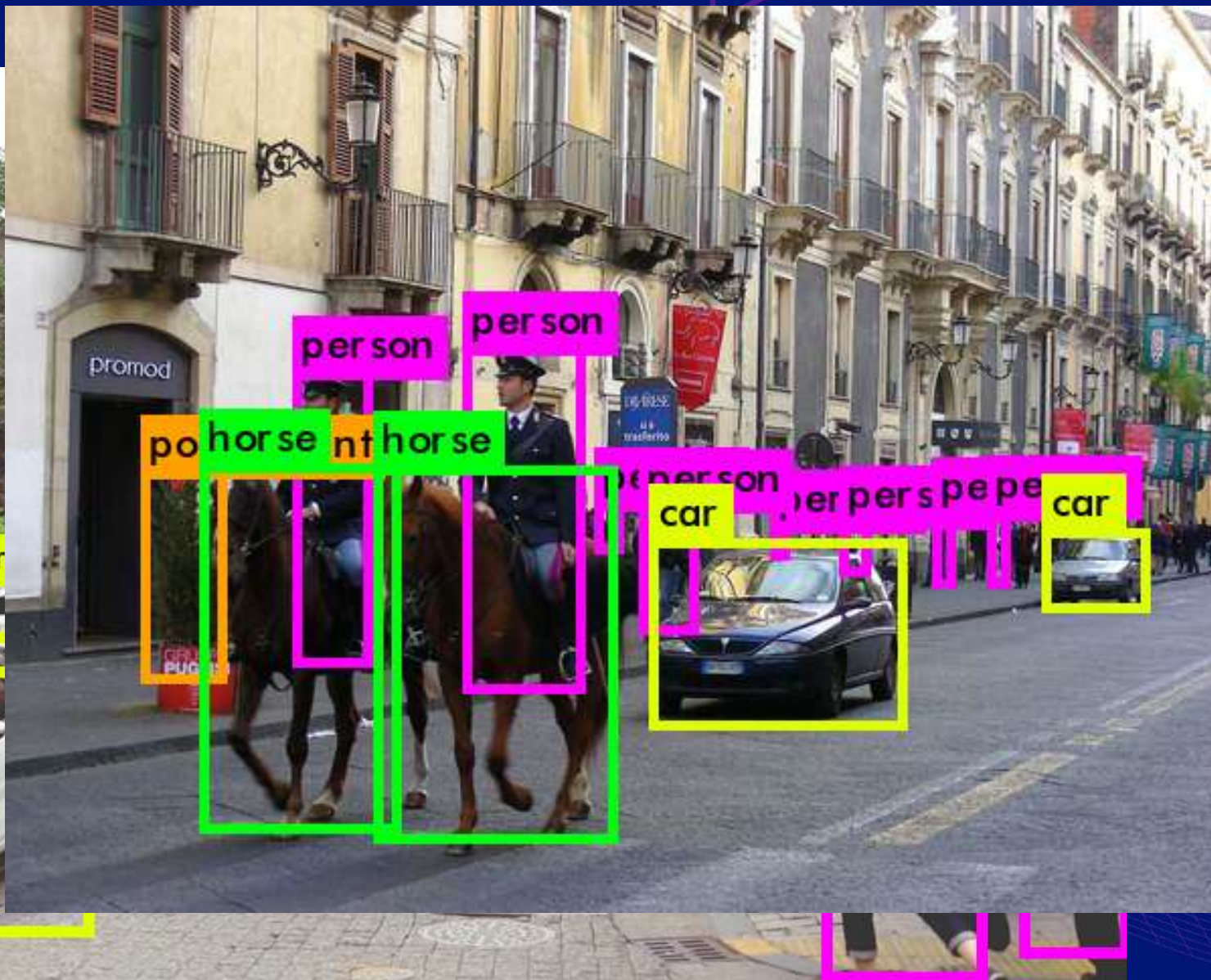


[Cheng et al., 2016]Wide& deep learning for recommender systems.

- ◆ Based on toured POI records, to predict user next interesting POIs.
- ◆ Using Time-LSTM model , to predict user behavior based on history log
- ◆ Equips LSTM with time gates to model time intervals. better captures both of users' short-term and long-term interests



[Zhu et al.,2016] What to Do Next: Modeling User Behaviors by Time-LSTM.In IJCAI,2017



AI探知：有品质/吸引力的产品

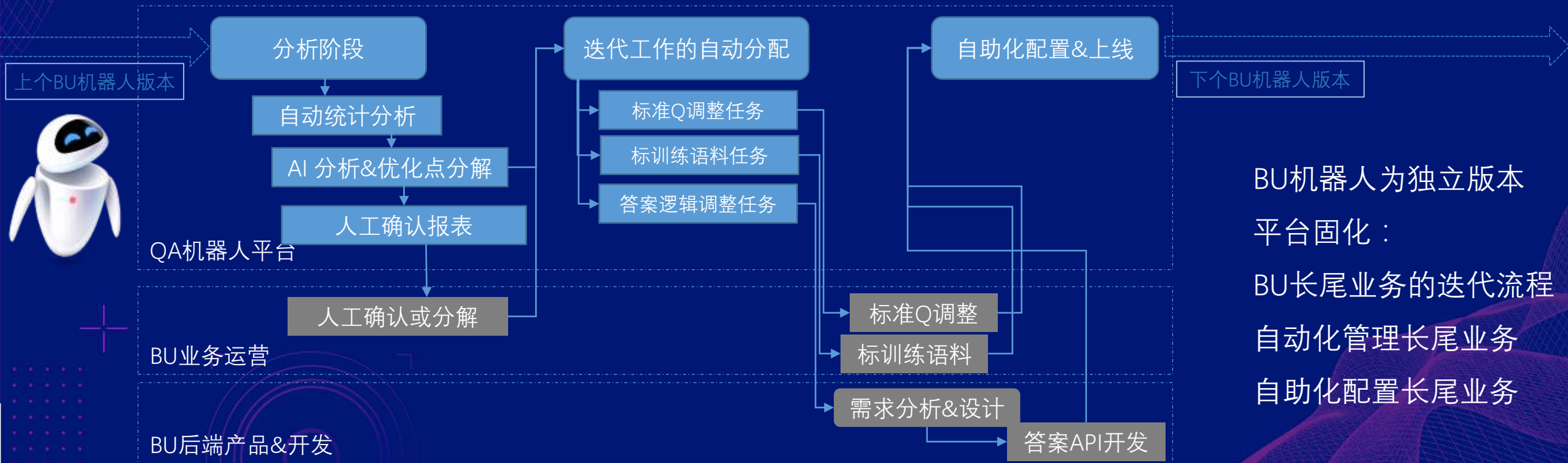
AI运营：有吸引力的产品信息

客服AI：有竞争力的产品服务

+ 基础AI能力的引入机制



公共AI团队：
平台的版本迭代
算法的版本迭代

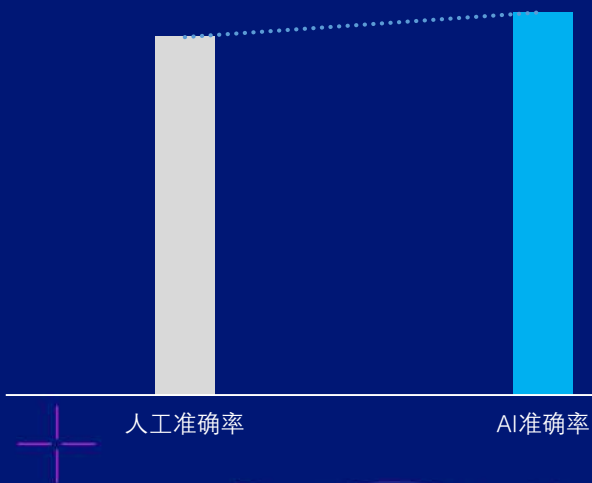


BU机器人为独立版本
平台固化：

BU长尾业务的迭代流程
自动化管理长尾业务
自助化配置长尾业务

准确率：AI>单人工标注

单轮会话，人工及AI识别准确率



自动分解

平台AI识别出
人工可能的误标注，

我的任务	我创建的任务	所有任务	新建任务
<div>待处理</div> <div>新增标准Q 截止日期 2018-10-11 03:33 陈挺 领取</div> <div>合并标准Q 截止日期 2018-10-11 03:31 陈挺 领取</div> <div>拆分标准Q 截止日期 2018-10-10 03:30 陈挺 领取</div> <div>消息标注</div>	<div>处理中</div> <div>修改标准Q 截止日期 2018-10-09 03:33 陈挺 去处理</div> <div>消息标注 截止日期 2018-10-17 03:24 陈挺 去处理</div> <div>删除标准Q 截止日期 2018-10-09 03:20 陈挺 去处理</div> <div>消息标注</div>	<div>待审核</div> <div>优化答案 截止日期 2018-10-18 03:25 陈挺</div> <div>新增标准Q 截止日期 2018-10-09 03:22 陈挺</div> <div>拆分标准Q 截止日期 2018-09-30 02:55 陈挺</div>	<div>已完成</div> <div>合并标准Q 截止日期 2018-10-09 03:19 陈挺</div> <div>修改标准Q 截止日期 2018-10-08 03:17 陈挺</div> <div>优化答案 截止日期 2018-10-09 03:16 陈挺</div> <div>拆分标准Q</div>

看板：明确的标注、抽检任务和迭代流程

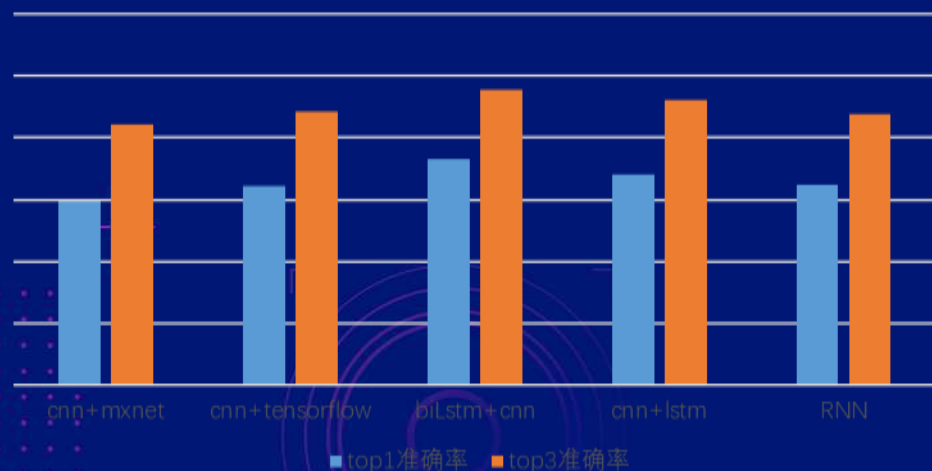


用自然语言，来完成交互式导购机器人

句子编码模型

- biLSTM可以从forward和backward两个方向获取上下文信息
- CNN模型对平移、缩放等具有不变形，可以较好检测出特征区域，鲁棒性好。
- 网络结构：pair-wise

算法对比



output layer

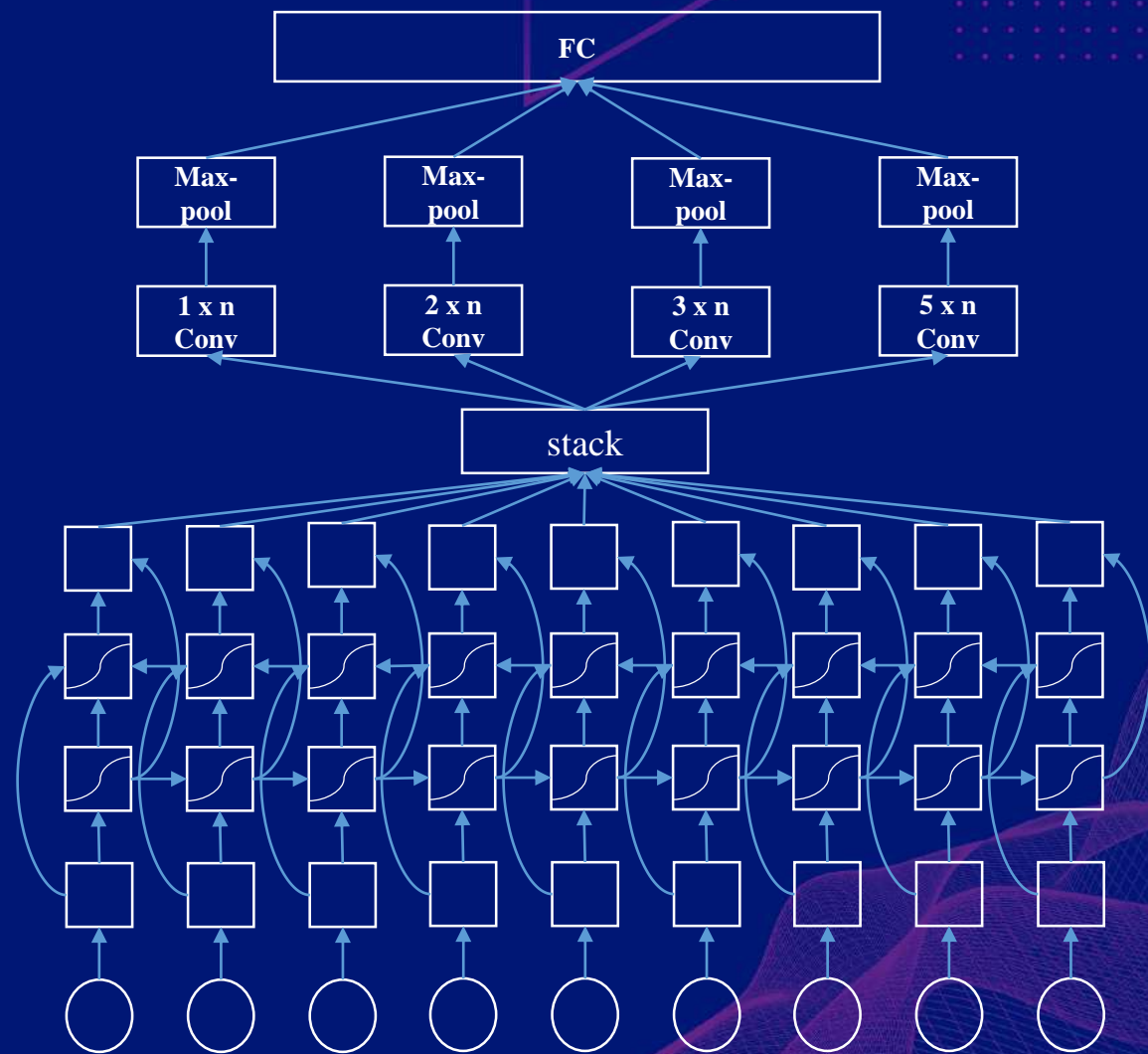
pooling layer

conv layer

biLSTM layer

embedding layer

input layer



基于BiLSTM-CNN的QA匹配

AI探知：有品质/吸引力的产品

AI运营：有吸引力的产品信息

客服AI：有竞争力的产品服务

基础AI能力的引入机制



AI基础能力引入：Ctrip Brain

Ctrip Brain 短域名：ai/

 **集团AI云平台**

首页

图像识别

图像审核

文字识别

人脸识别

数字水印

语音识别

SDK文档

数据运营

退出

动物识别

植物识别

图像主检测

车辆识别

Logo商标识别

菜品识别

图像标签识别

看图说话

场景物体识别

相同图检索

相同图入库

相同图删除

语音识别

了解更多

 新品上线

人脸识别服务，全新上线，欢迎体验

 解决方案

智能售前导购服务，打开售前新视界

 最新论文

Learning from Simulated and unsupervised Images through...

 众筹标注

按条付费的标注体系，赚点零花钱

➤ 图像识别

- 动物识别
- 植物识别
- 图像主检测(百度&腾讯)
- 车辆识别
- Logo商标识别
- 菜品识别
- 图像标签识别(百度&腾讯)
- 看图说话
- 地标识别
- 相似图识别(百度&腾讯)
- 图像水印识别

➤ 图像审核

- 广告审核
- 色情审核
- 政治敏感审核
- 暴恐审核
- 恶心图审核
- 头像审核
- GIF审核

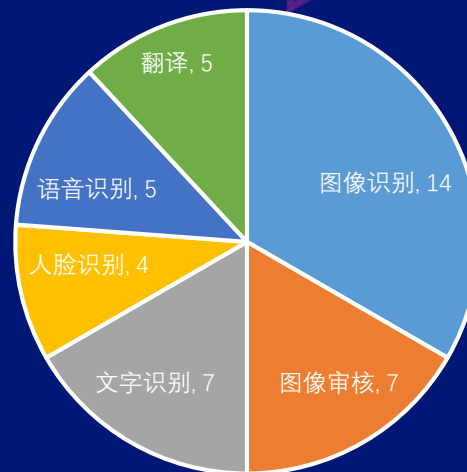
➤ 文字识别

- 证件全文本OCR
- 通用文字识别(百度&腾讯)
- 身份证识别
- 驾照识别
- 行驶证识别
- 银行卡识别

➤ 人脸识别

- 人脸识别(百度&腾讯)
- 人脸验证
- 活体检测

AI能力分布



同类服务测评

阿里录音文件识别

录音文件识别:

“录音文件识别”服务是以RESTful方式提供长语音文件识别接口。基于该Restful API接口，开发者可以方便获取语音识别能力。

功能介绍:

- 支持两种调用方式：轮询方式和回调方式
- 支持单轨/双轨的wav格式、MP3格式的语音文件识别
- 目前只支持8k和16k的采样率；
- 支持汉语普通话，欧美英语识别

测评报告:

音频	音频文件大小	人工标注字符数	识别错误字符数	准确率
286278704.wav.1	7.46M	927	0	100%
286278704.wav.2	7.46M	1224	0	100%
286570173.wav.1	6.63M	942	0	100%
286570173.wav.2	6.63M	1508	0	100%
286633600.wav.1	4.66M	738	0	100%
286633600.wav.2	4.66M	455	0	100%
286683374.wav.1	7.24M	1627	0	100%
286683374.wav.2	7.24M	1174	0	100%

腾讯长语音识别

功能介绍:

长语音识别依托腾讯的语音技术，为开发者提供全面优质的近场长段语音转文字功能。目前支持汉语普通话，并涵盖南北地域重口音。除了常规语音识别技术之外，还基于长语音中更多的语义信息，对语音识别结果进行更精细的纠错、智能断句、格式转换等处理，识别结果更精确。

测评报告:

音频	音频文件大小	人工标注字符数	识别错误字符数	准确率
286278704.wav.1	7.46M	927	0	100%
286278704.wav.2	7.46M	1224	0	100%
286570173.wav.1	6.63M	942	0	100%
286570173.wav.2	6.63M	1508	0	100%
286633600.wav.1	4.66M	738	0	100%
286633600.wav.2	4.66M	455	0	100%
286683374.wav.1	7.24M	1627	0	100%
286683374.wav.2	7.24M	1174	0	100%



AI基础能力引入：Ctrip Brain

Ctrip Brain使用情况：

日均调用超过**千万次**/天，外部AI赋能占比达到**72%**

本PPT来自2018携程技术峰会
更多技术干货，请关注“携程技术中心”微信公众号

