



携程技术中心

# 携程技术沙龙

## 携程AI平台及客服机器人

分享人：于磊

# 目录

CONTENTS

- 1 携程AI—Democratize AI
- 2 携程客服机器人&知识图谱
- 3 携程小诗机
- 4 Q&A

# Democratize AI

合作、互补、共享 ➡ 共赢、高速创新 ➡ 下一代的OTA用户体验

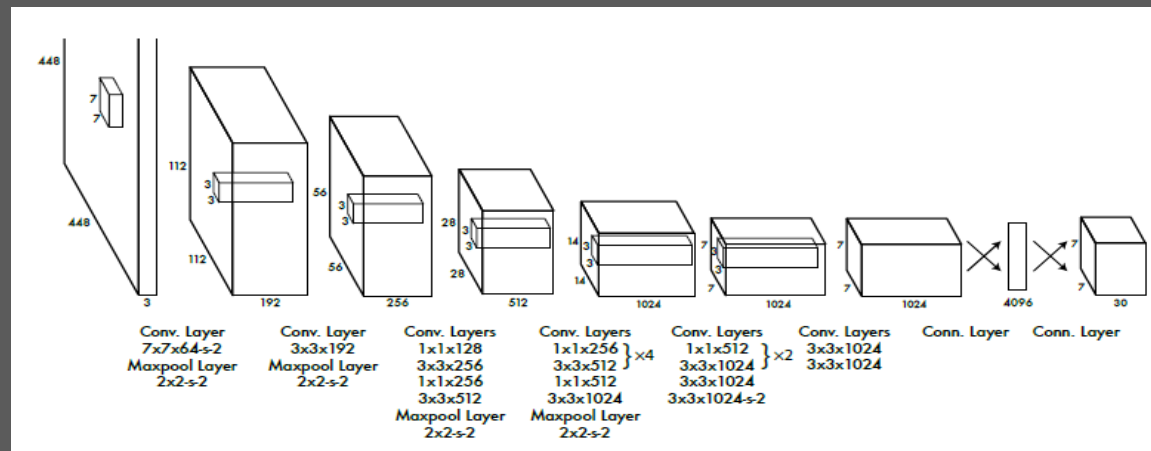




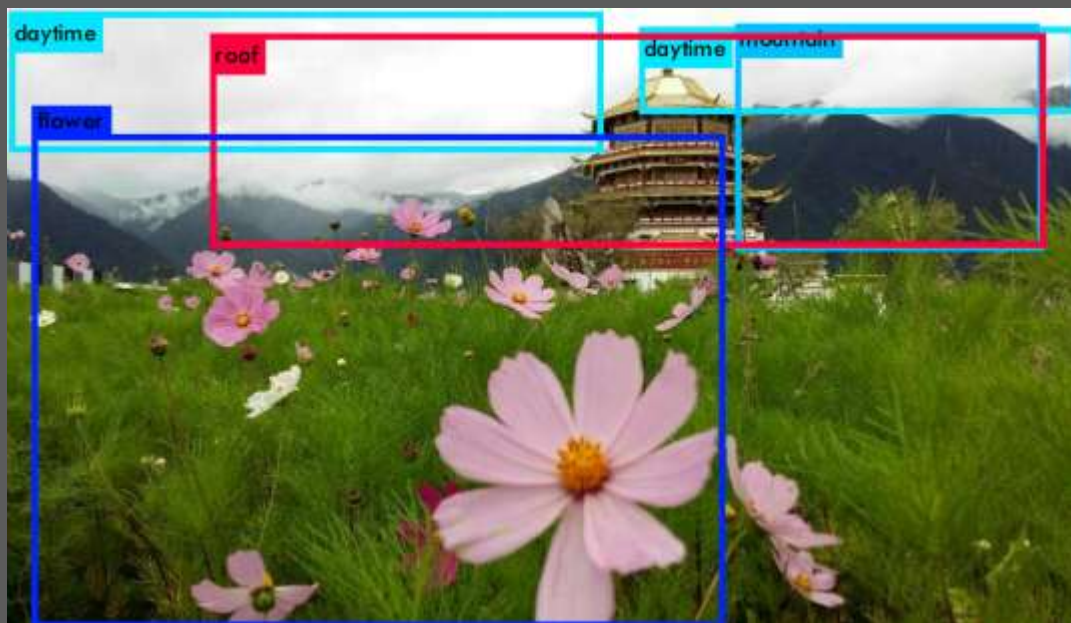


## 图片检测(Image Detection)

- 背景：检测图片物体标签，作为素材提供给写诗引擎
- 语料：类别120，训练语料10万条
- 24层卷积层+2层全连接层

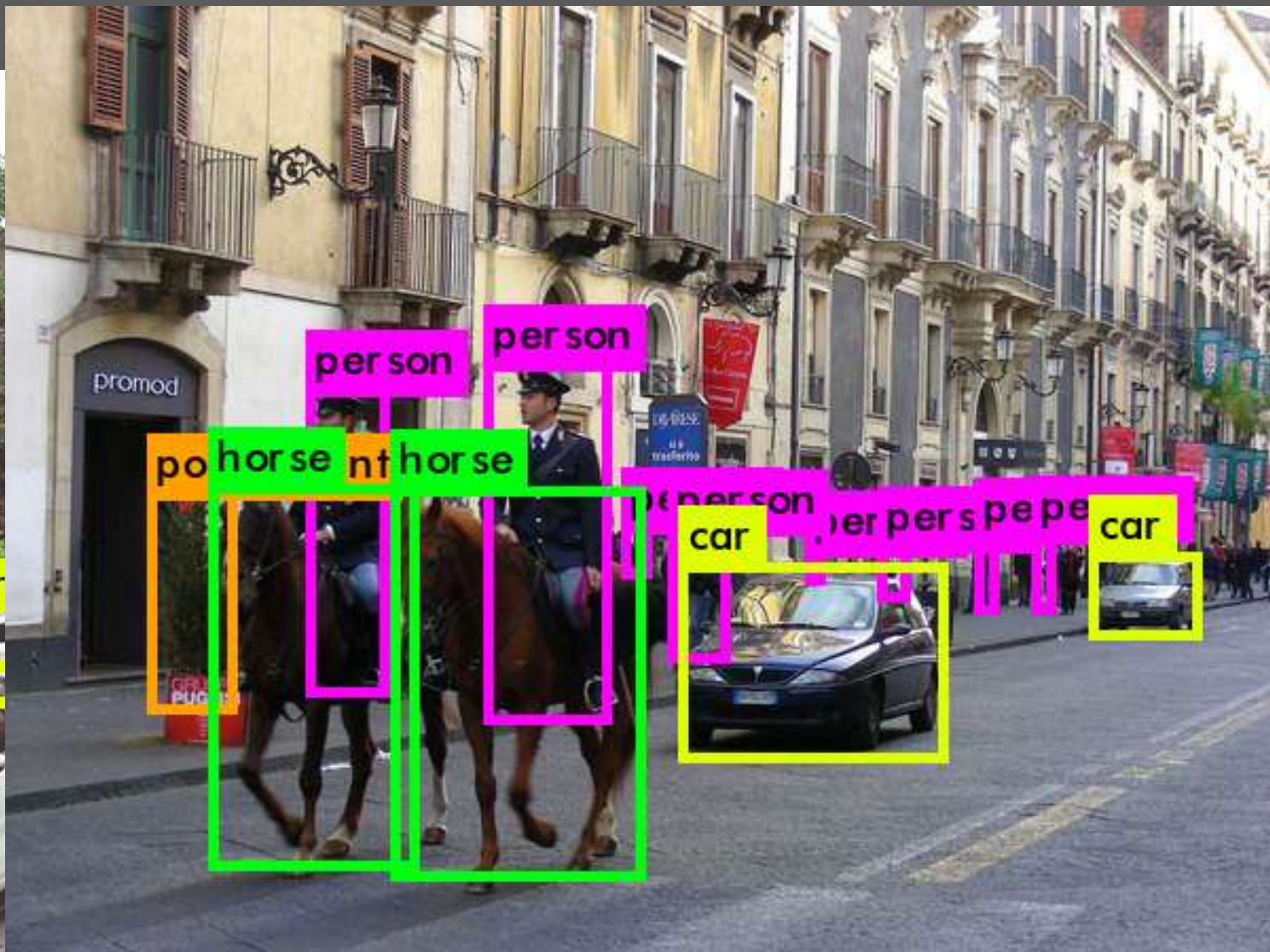


Yolo网络结构





## 图片检测(Image Detection)





## 视觉图像





## 智能个性化



## 智能导购





# 目录

## CONTENTS

1

携程AI平台

2

携程智能客服&知识图谱

3

携程小诗机

4

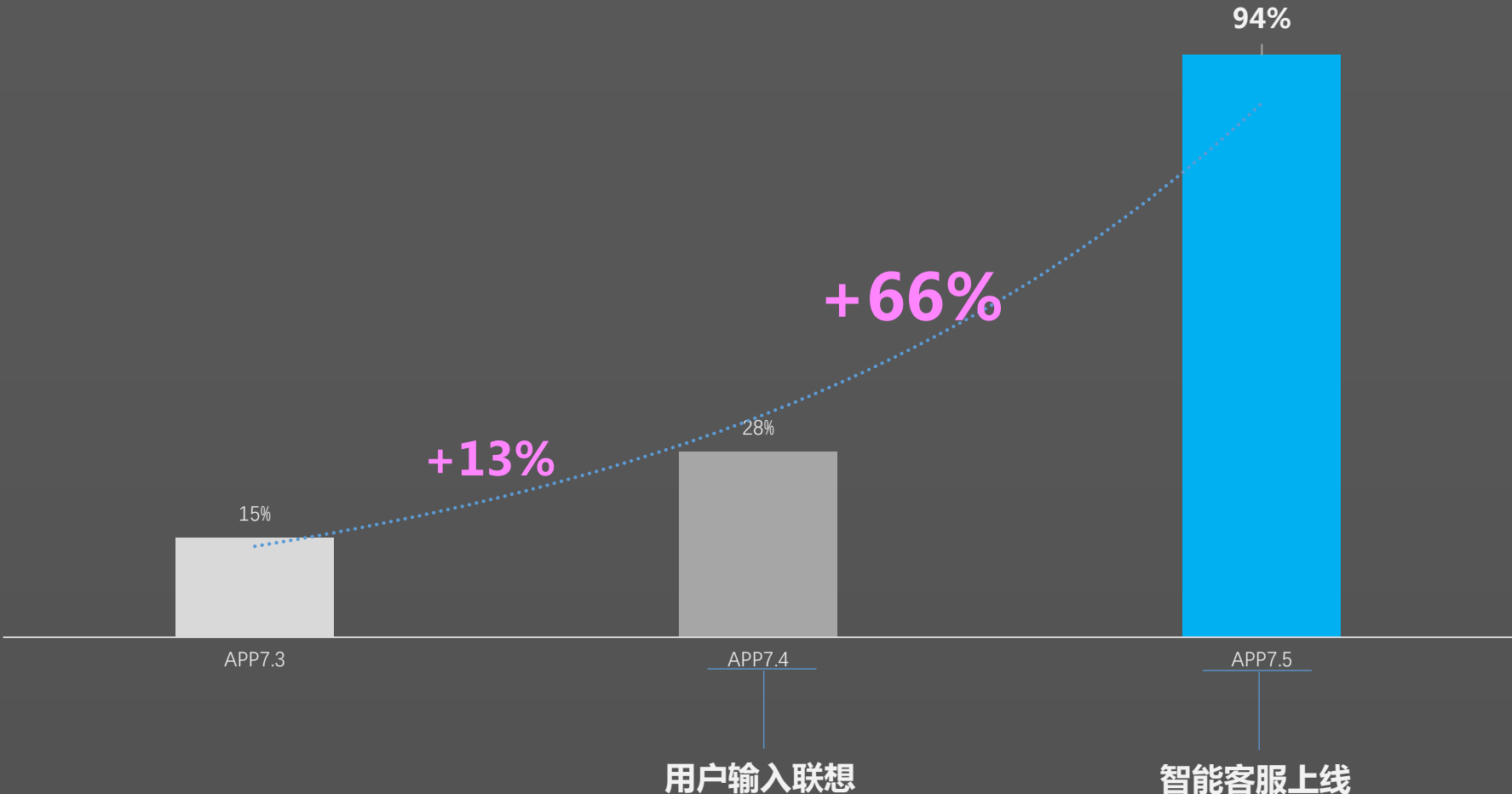
Q&A

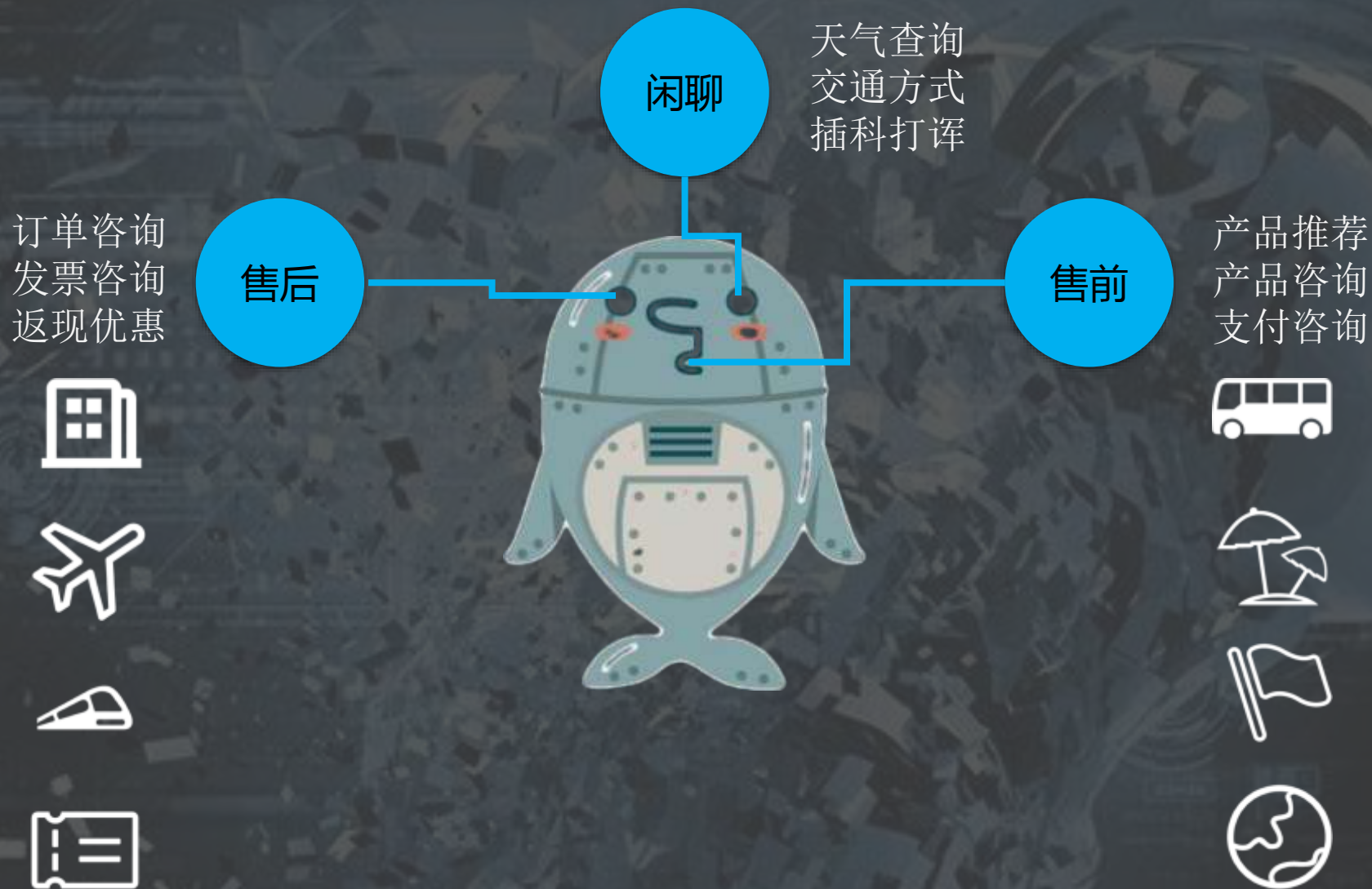


智能客服



酒店客服可自助场景解决率





## 应用架构



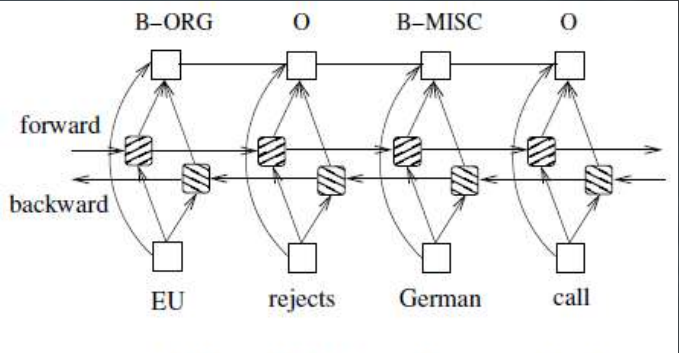


## 命名实体识别

- 应用场景
  - ✓ 售前售后景点、目的地等识别
  - ✓ 旅游名词识别
  - ✓ KB-QA中识别主题信息

### 常用模型

Model	优势
RNN	获取长远的上下文信息，拟合非线性能力
CRF	考虑整个句子的局部特征的线性加权组合，目标是 $p(y x;w)$ 使得序列的联合概率最大
HMM	基于马尔科夫性质
LSTM+CRF <sup>[1]</sup>	利用两个模型的优势



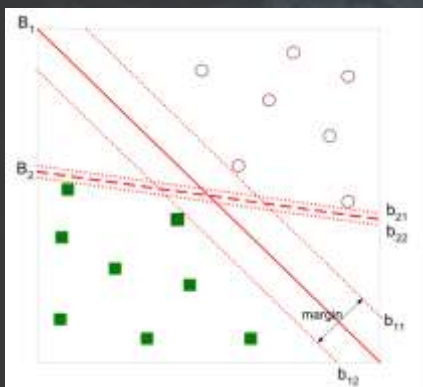
```
我想下个月5号去杭州看西湖
ParseResult:
{
  u'City': {
    u'start': [],
    u'unknown': [],
    u'dest': [{u'type': u'city', u'id': 17, u'name': u'杭州'}]
  },
  u'IsDiscount': -1,
  u'error': u'',
  u'Poi': [{u'id': 91557, u'name': u'西湖'}],
  u'Date': {
    u'StartDate': u'2017-10-5',
    u'EndDate': u'',
    u'Days': u'',
    u'UnknownDate': u''
  },
  u'TripType': [],
  u'Level': -1,
  u'TopicLabel': []
}
```

原句：我要开发票：衢州恒业汽车部件有限公司，税务识别号9133080275590610X5。邮寄地址：浙江省衢州市柯城区三江东路2号卢森堡A座浙江青风律师事务所戴剑波收，手机13305703919  
 解析：我要开发票：[institution]，税务识别号9133080275590610X5。邮寄地址：[address][person]收，手机[mobile]

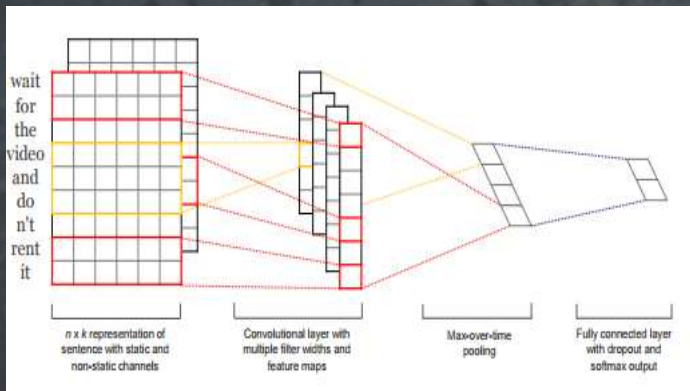
[1]. Z Huang, W Xu, K Yu. Bidirectional LSTM-CRF Models for Sequence Tagging - arXiv preprint arXiv:1508.01991, 2015

## 用户意图识别

- 基于用户问题，使用意图分类模型进行用户意图的判定，从而将用户问题定位到某个子领域。
- 辅助进行进一步的用户语义解析
- 基于意图信息，进行用户对话管理和状态分析

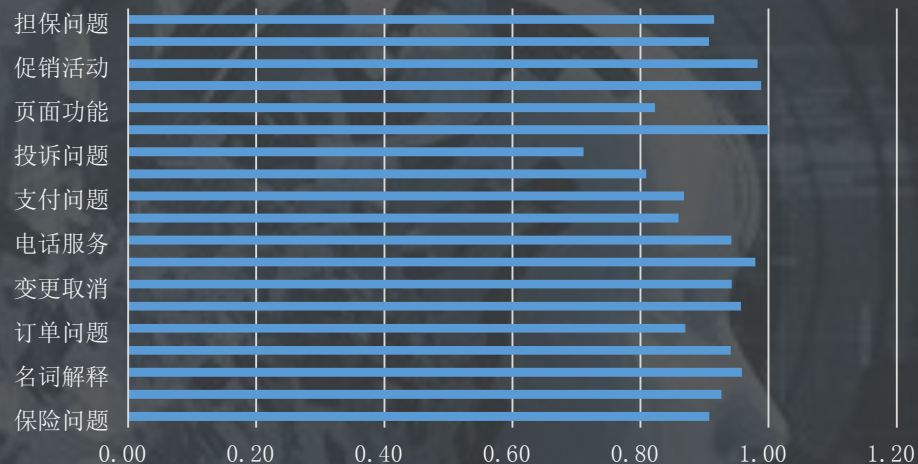


SVM分类模型

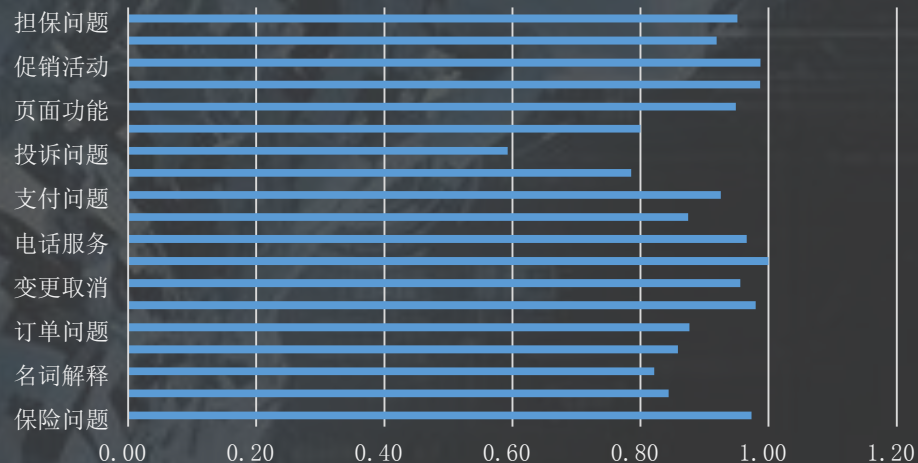


CNN分类模型

Precision



Recall

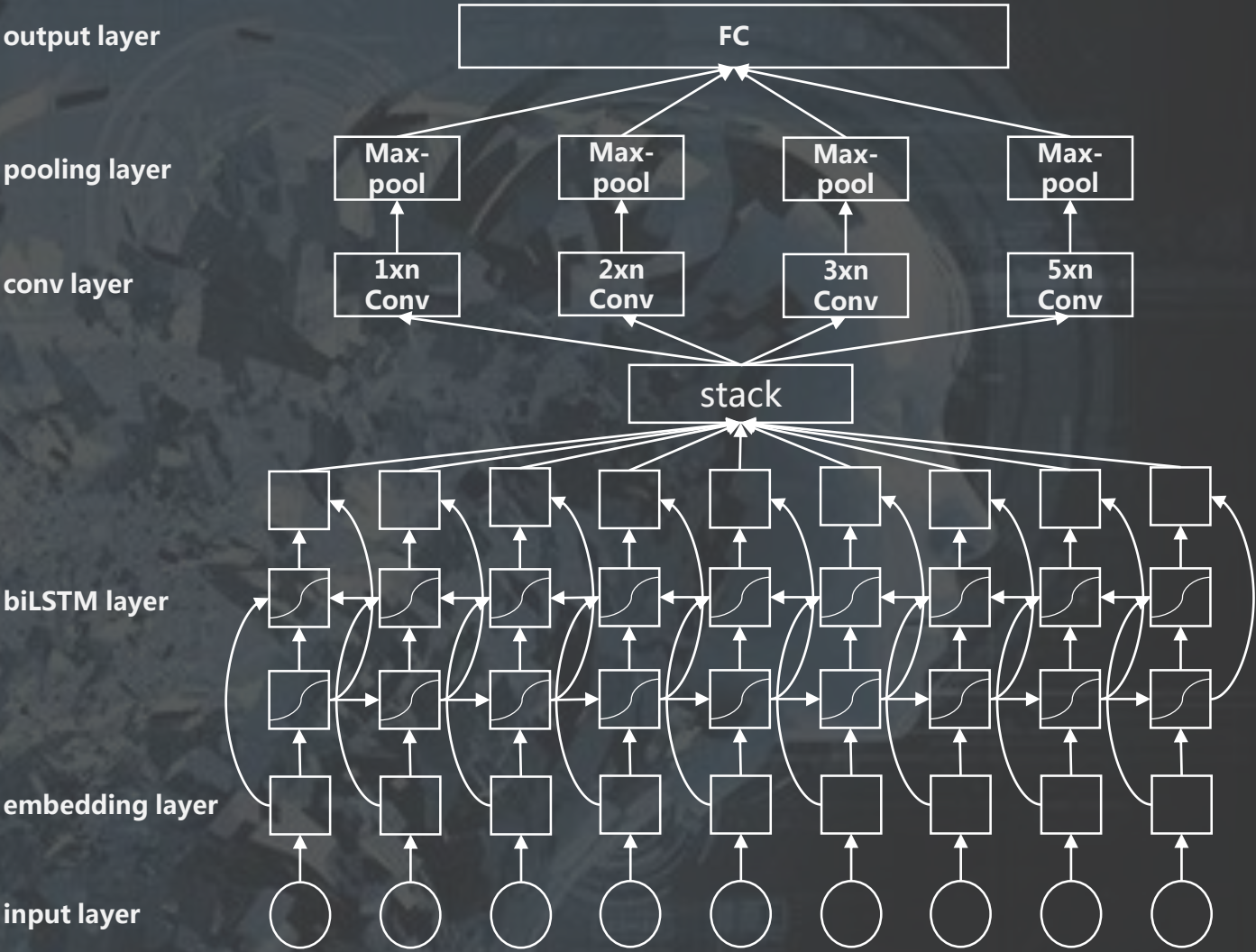
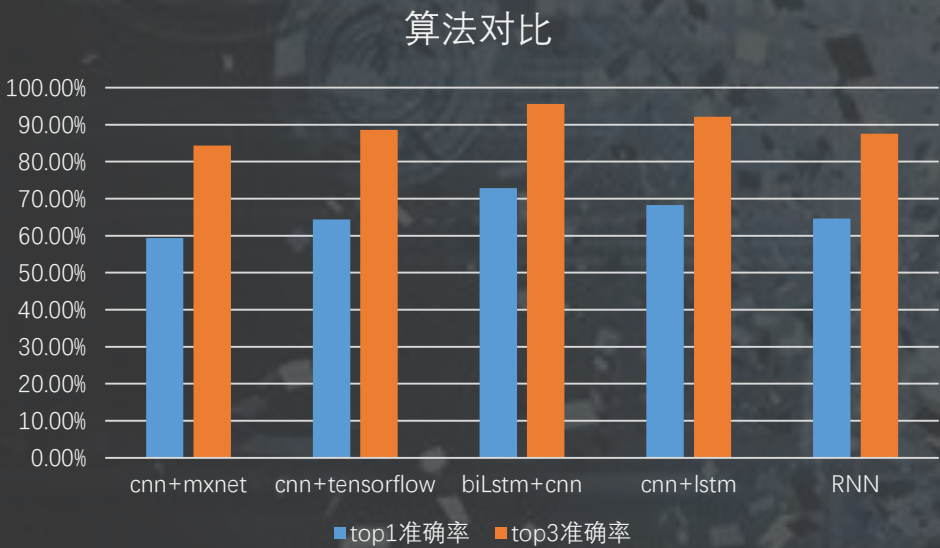




## 标准Q匹配

### 优势

- 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。
- 最终，通过多组实验比较，发现biLSTM+CNN离线效果最佳。



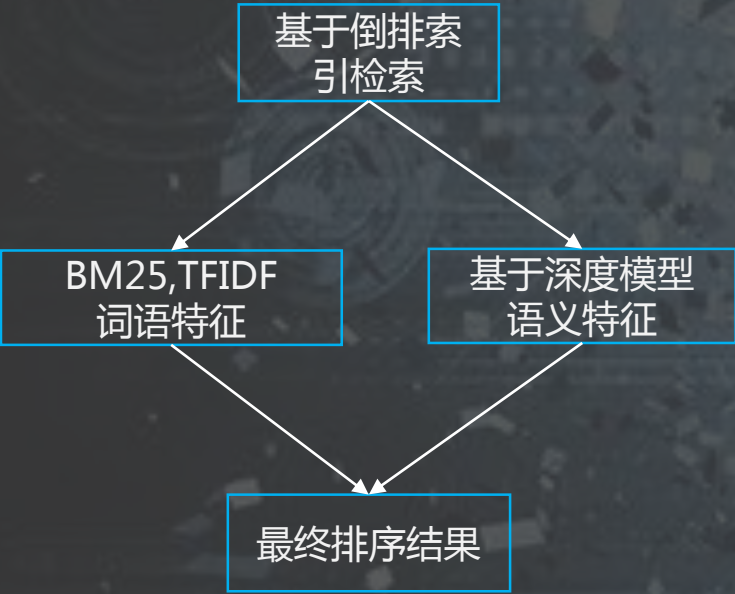
基于biLSTM-CNN的QA匹配

[1] Applying Deep Learning to Answer Selection: A Study and An Open Task Minwei Feng, Bing Xiang, Michael R. Glass, Lidan Wang, Bowen Zhou ASRU 2015  
 [2] Lai S, Xu L, Liu K, et al. Recurrent Convolutional Neural Networks for Text Classification[C]//AAAI. 2015, 333: 2267-2273.  
 [3] Wen Y, Zhang W, Luo R, et al. Learning text representation using recurrent convolutional neural network with highway layers[J]. arXiv preprint arXiv:1606.06905, 2016.



## QA-检索

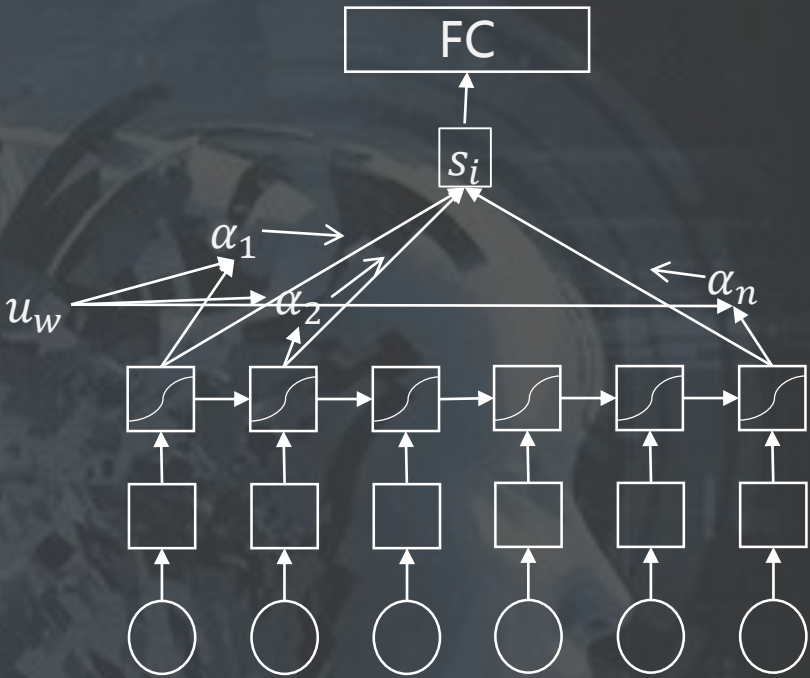
- 类似个性化推荐，分两步：检索+排序
- 适用QA量大的场景，例如闲聊，问答检索等
- 基于attention机制可以提升QA关联度



$$u_{it} = \tanh(W_w h_{it} + b_w)$$

$$\alpha_{it} = \frac{\exp(u_{it}^T u_w)}{\sum_t \exp(u_{it}^T u_w)}$$

$$s_i = \sum_t \alpha_{it} h_{it}$$



基于检索的QA

>>北京的景点有什么

北京旅游景点之一：北海公园：位于故宫西北角，始建于十世纪初，是中国现存最完整的古代园林之一，公园面积78公顷，其中水面有39公顷。

北京旅游景点之二：大钟寺。原名觉生寺，座落在北京城的北三环西路，建于1733年。因寺内藏有明代永乐年间铸的大钟，俗称大钟寺，永乐大钟高近7米，外径3.3米，重达46.5吨；钟身铸满佛教经咒，约23万字。钟声清朗，方圆百里可闻。

北京旅游景点之三：天坛。明、清两代帝王祭天祈谷的地方，建于1420年。坛墙南方北圆，象征“天圆地方”。

北京旅游景点之四：十三陵。位于八达岭长城东南不远的大寿山下，这里埋葬着明代的13个皇帝。从1409年始建长陵起，到1644年修建思陵止，历时200余年。1956年，中国考古工作者发掘出万历皇帝朱翊钧的陵墓——定陵，发现了地下宫殿，除棺槨外，还出土了3000余件珍宝。北京旅游景点之五：雍和宫。在北京城区东北角，建于1694年，原为清第三位皇帝雍正皇帝的府邸，1744年改建为喇嘛庙。庙中藏有18米高的白檀木大佛，由一整棵檀木雕成。

>>上海到杭州怎么走方便

动车

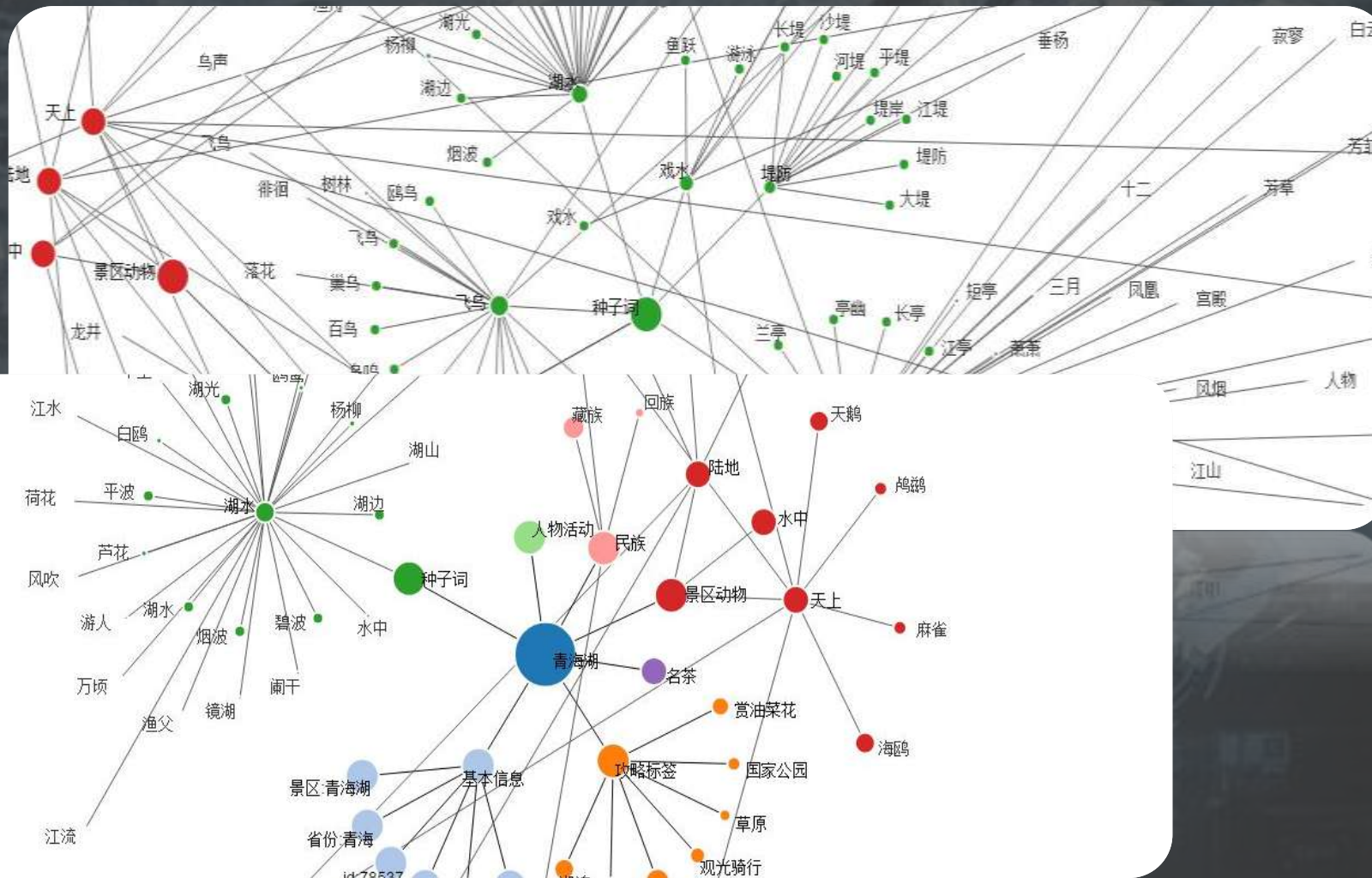
>>我要去内蒙古，要注意什么

内蒙古旅游详情：<http://you.ctrip.com/place/neimenggu100062.html> 参考一下

[1] Lowe R, Pow N, Serban I, et al. The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems[J]. arXiv preprint arXiv:1506.08909, 2015.

[2] Yang Z, Yang D, Dyer C, et al. Hierarchical Attention Networks for Document Classification[C]//HLT-NAACL. 2016: 1480-1489.

## 知识图谱KB-QA

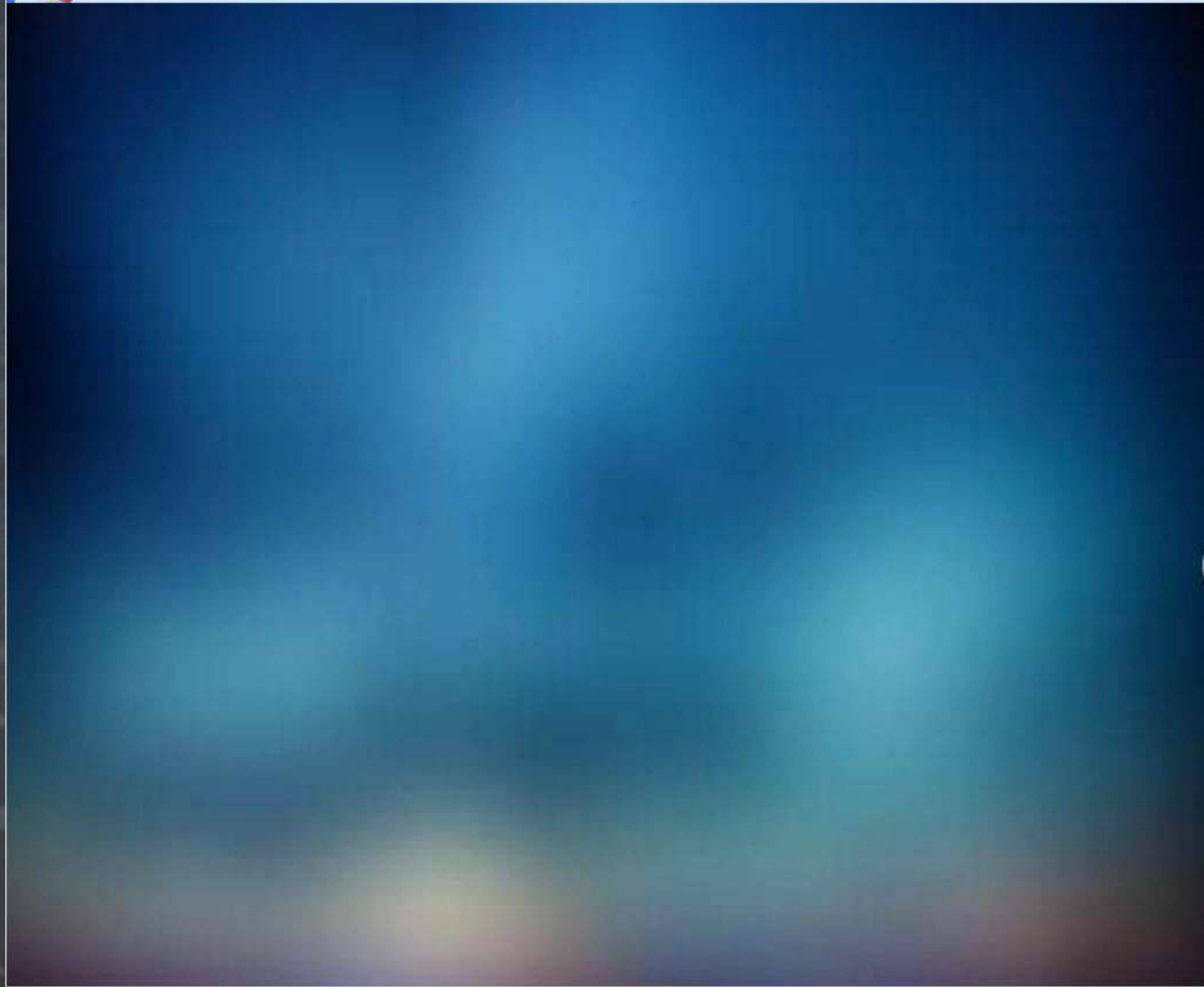




售前智能导购



携程客服机器人



✓ 发送



# 目录

## CONTENTS

- 1 携程AI平台
- 2 携程客服机器人
- 3 携程小诗机
- 4 Q&A

## Input:

景点写诗

城市写诗

天气、季节

心情

照片写诗

人名写诗

## 格律：

宝塔诗

律诗

绝句

景物识别&  
成诗：  
红楼；  
梅花；

性别及表情  
识别：  
严肃；  
青年男性。





## Input:

景点写诗

城市写诗

天气、季节

心情

照片写诗

人名写诗

## 格律：

宝塔诗

律诗

绝句

照片地点  
知识库：  
西湖；  
西湖的知  
识库



时间写诗：  
秋；  
傍晚；  
天气知识库；



郭德纲  
老郭诙谐通道术，  
融和春意动剑眉。  
大话天仙妙绝伦，  
喜嗔语默皆相宜。

人物画  
像：  
职业；  
作品；  
身材；风  
格



基于情境的感  
情：  
春节前最后一  
天；



人机盲测 第3名(11场)

与上海地区诗人比，略胜

	冠军	亚军	季军
赵玉玮	5	0	1
邹竞夫	2	6	2
AI LEE	2	3	2
卿云子	2	2	2
沈以昕	0	0	2
张一双	0	0	0

	思念	孤独	夏之愉悦	秋之悲伤	北京的雾	杭州的雨	黄山的雪	青海湖的晴	沙-正宝塔	沙-倒宝塔	景点 愉悦
赵玉玮	7.45	7.155	6.72	6.47		6.52	6.73			6.68	6.72
邹竞夫	8.37	7.575	6.705	6.49	8.22	6.41	6.685	6.32	6.425	6.24	6.375
AI LEE	7.885	8.14	6.195	6.535	7.61	6.225	6.07	6.4	6.15	5.91	6.4
卿云子		8.395	6.315		8.655		6.6			6.365	6.515
沈以昕	7.725	6.725	5.85	6.32	7.565	6.395			5.61		6.31
张一双	7.115			6.155			6.325	5.87	5.175		6.115

目前 远超竞品

宝塔寺、情感、季节、天气、人物、绝句、律诗、城市、美食

西湖	西湖	西湖. 夏
凉风乱入杭城路， 秋水孤山似旧年。 遍地桂花停过客， 侵帘绿影满湖边。	绿树余杭花草畔， 通都月色谁自怜。 喜听潇雨鸣枫叶， 安得凉风起桂筵。 久靠林荫堤岸伫， 长依湿地碧波边。 西湖桥下潺湲水， 凭有孤山却不前。	重云深处引溪涧， 湖水孤山路远长。 遥爱断桥残夏早， 风清杨柳映通方。 莲花倚槛雨生雾， 荷叶绕船橹走旁。 林缝落光波潋潋， 寺边湿地色苍茫。

西湖. 雪  
城巷夜空却白昼，  
西湖山色故依然。  
敢忘湿地凌冬约，  
不负君梅岁竹缘。  
万里重云寒水边，  
满城娑树雪中天。  
萧疏枫叶扁舟岸，  
灵隐冷风谁顾怜。

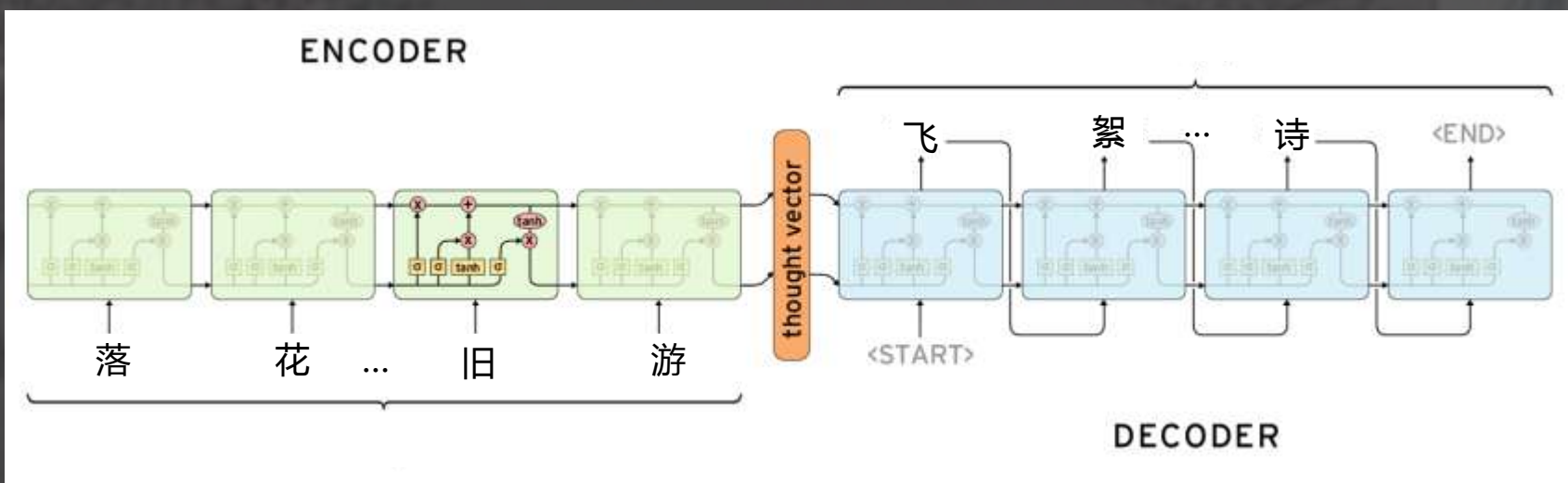
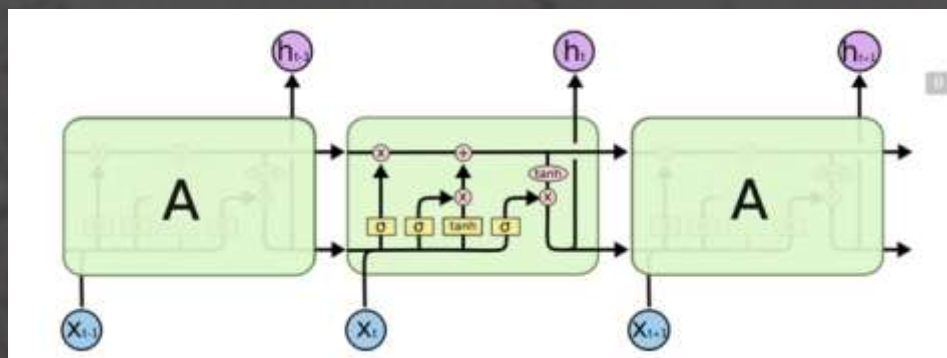
澜	香蕊燕雀凉醉颊
惊湍	玉屑梅蕊翠靥
水流漫	偷香团扇妾
鲸波海观	芳情细贴
江涌万卷蟠	花重叠
涛漱掀翻浪寒	春叶
激荡澎湃渺云端	蝶





## 古诗自动生成

基于LSTM模型和seq2seq相结合，进行诗歌的创作



春日

落花飘红忆旧游,  
飞絮皱绿赋新诗.  
江风吹雨又黄昏,  
望月连云消白日.



携程技术中心

# THANK YOU!

## Q&A



小诗机——一键成诗



更多技术干货可关注  
“携程技术中心” 微信公号