

ArchData

技术峰会北京站

主办方：



2017年9月24日北京海淀区丹棱街5号微软亚太研发中心一号楼一层 故宫会议室

f

深度学习与智能对话机器人

吴金龙@爱因互动

吴金龙

- 2005~2010：北大数学院
 - 推荐算法
- 2010~2011：阿里云
 - PC/手机输入法
- 2011~2017：世纪佳缘
 - 用户推荐、网警等数据系统
 - 技术部负责人
 - 一个AI负责人
- 2017~现在：爱因互动
 - 技术合伙人
 - 算法负责人
- **ChatbotsChina**发起人
 - bot 相关的技术、产品、运营
 - 微信公众号/交流群、微博
 - 线下活动

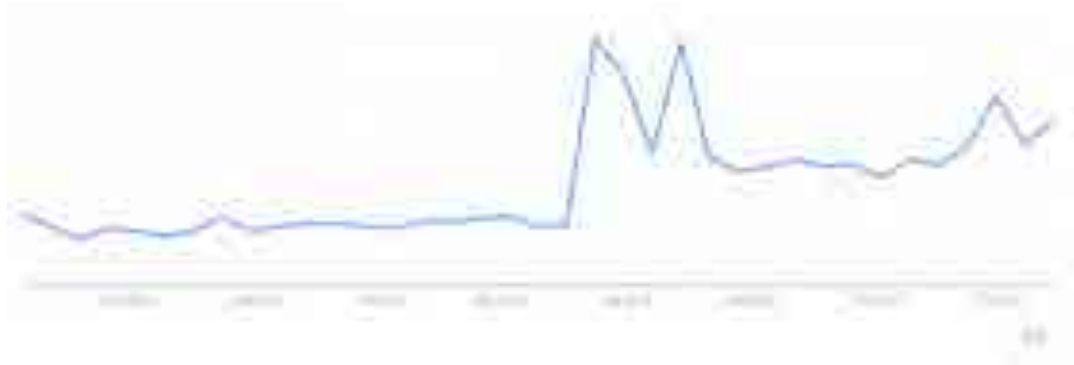
目录

- 对话机器人的兴起
- 爱因互动的Bot技术：DeepBot及其他
- 对话机器人的商业实践

对话机器人简史



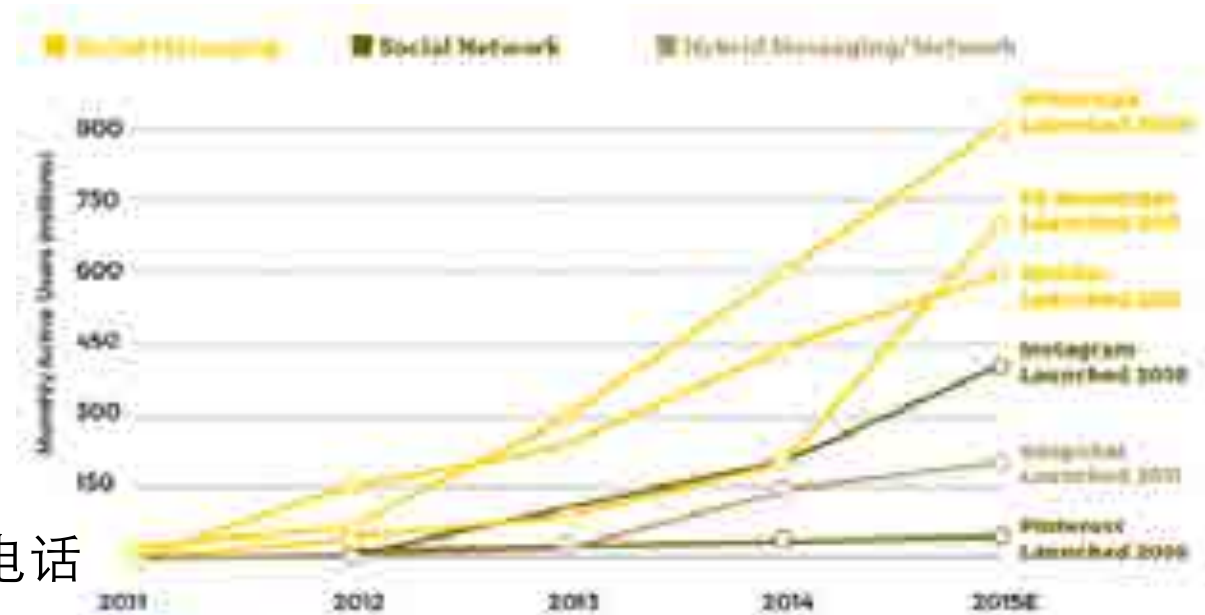
对话服务爆炸式增长



Chatbot 在 Google Trends 中搜索量的变化

- 2016.3 微软 Bot Framework
- 2016.4 Facebook Messenger 平台
- 2016.5 Google Assistant
- 2016.6 苹果 Siri SDK

- APP 下载量比前一年下降20%
- 前3位 APP 占据80%时间
- 过去2年消息平台爆发式增长
- 70%用户倾向于文本消息而非电话



主流消息平台使用增长情况

对话产业

Designed by
Dan Miller, Derek Top &
Nicolas De Kouchkovsky

Intelligent Assistance landscape

powered by
VB | Profiles



对话机器人未来趋势

对话机器人正在被广泛应用

以对话的方式提供服务，完成用户指令

在未来3-5年将成主流的人机交互方式，并由此带来广阔的市场前景



个人信息助理

信息整理收集

行程/会议安排

写作助手



客服/导购机器人

信贷员/理财师

3C产品导购

化妆品导购

保健品导购



泛娱乐/教育聊天机器人

车载娱乐

社交陪伴

初级技能传授

爱因互动 EIN+

- 为企业提供人工智能对话解决方案
- **C**onversation **a**s **a** **S**ervice



行业解决方案

深度结合行业相关知识和领域信息，
为重点客户定制专属解决方案。



对话交互 (CUI) 规范指导

以自然的、智能的多轮对话作为人机交互的统一界面，
免去用户学习成本。



简单优雅的接入实施

提供 RESTful API，微信公众号和 Web SDK 接入方式，
覆盖多平台使用。



管理控制台

在这里查看数据报表，查看聊天日志，
自动提交语料，获取每月账单，管理 API Token。

爱因客户

vanke 万科

宜信 CreditEase

网信

盈米财富

途虎养车
TUHU.CN

G7 智慧物联网

Glodon 广联达

订购小助手

keep

Westone 万士通

保隆

菲菱说

合作伙伴

QINGCLOUD 青云

美洽

DaoCloud

环信

NVIDIA

合作示例

- 订餐小秘书

The screenshot shows a chat interface for '订餐小秘书' (Ordering Assistant). The interface is divided into three main sections: a chat area on the left, a user profile section in the middle, and an order management section on the right. The chat area shows a conversation where the user asks for a recommendation, and the assistant responds with suggestions. The user profile section displays basic information such as name, phone number, and address. The order management section shows a list of orders with a '订餐' (Order) button.

Annotations on the left side of the screenshot point to specific features:

- 槽位提取** (Slot Extraction): Points to the user's input text in the chat.
- 意图识别** (Intent Recognition): Points to the assistant's response text.
- 多轮对话** (Multi-turn Dialogue): Points to the entire chat history.
- 回复生成** (Response Generation): Points to the assistant's response text.
- 商品推荐** (Product Recommendation): Points to the assistant's response text.

爱因机器人类型



ASK MORE
QUESTIONS

问答机器人
回答常见用户问题

推荐机器人

通过几个问题获得用户需求推荐产品



信息采集机器人



任务机器人

通过多轮与上下文记忆完成特定任务



闲聊机器人

随便聊两句呗

多通道API



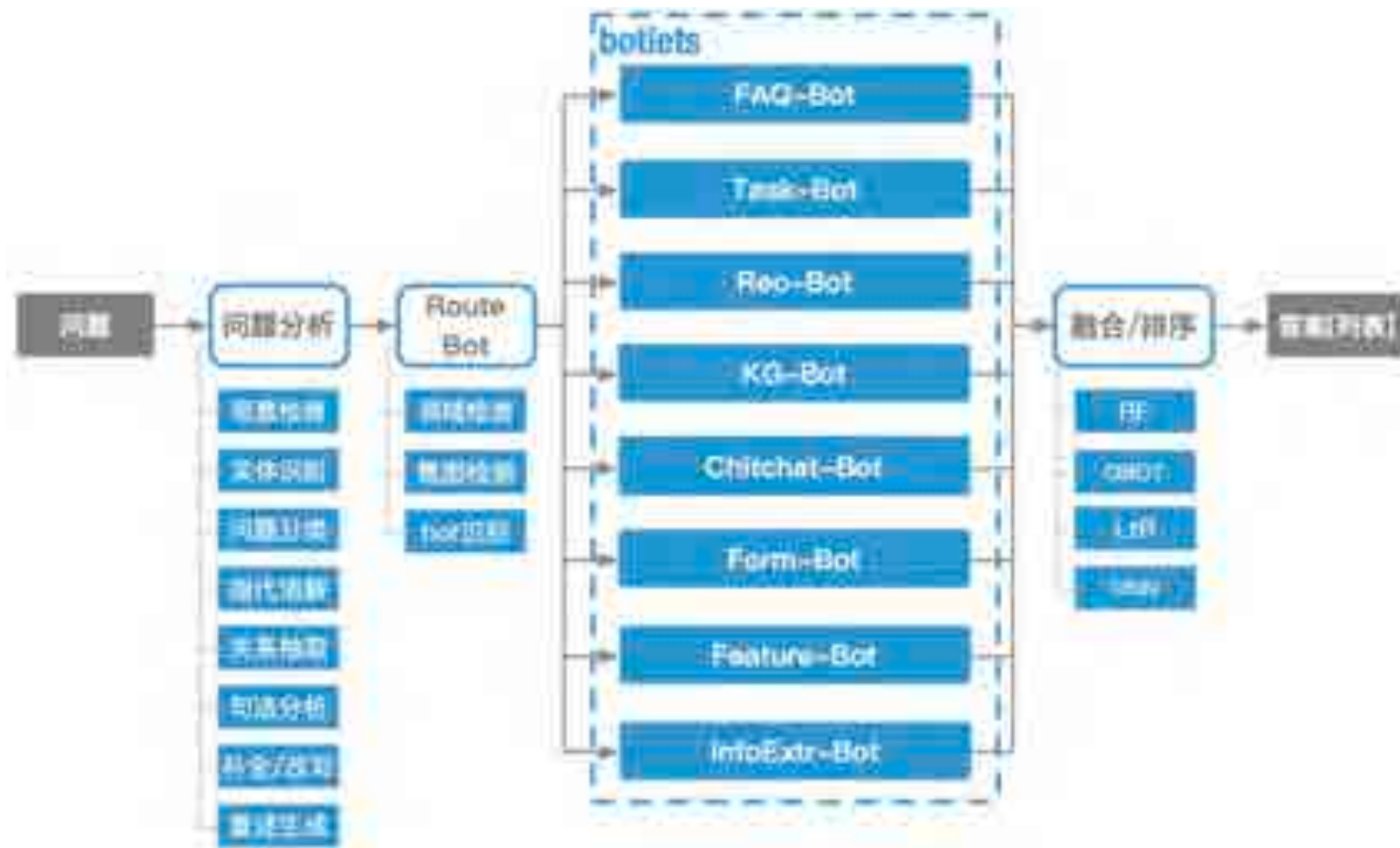
高可用工程框架



DeepBot对话框架

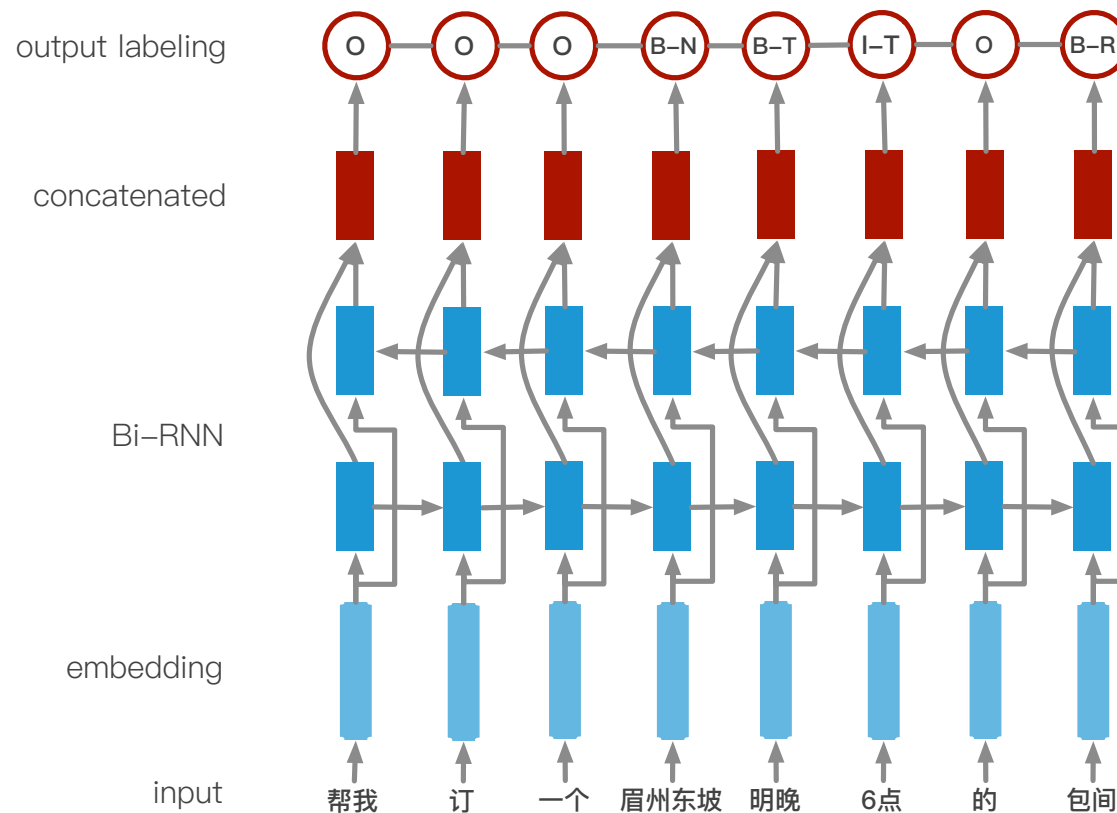
- 融合

- 不同的问题,
- 不同的框架,
- 不同的模型



问题分析——实体识别

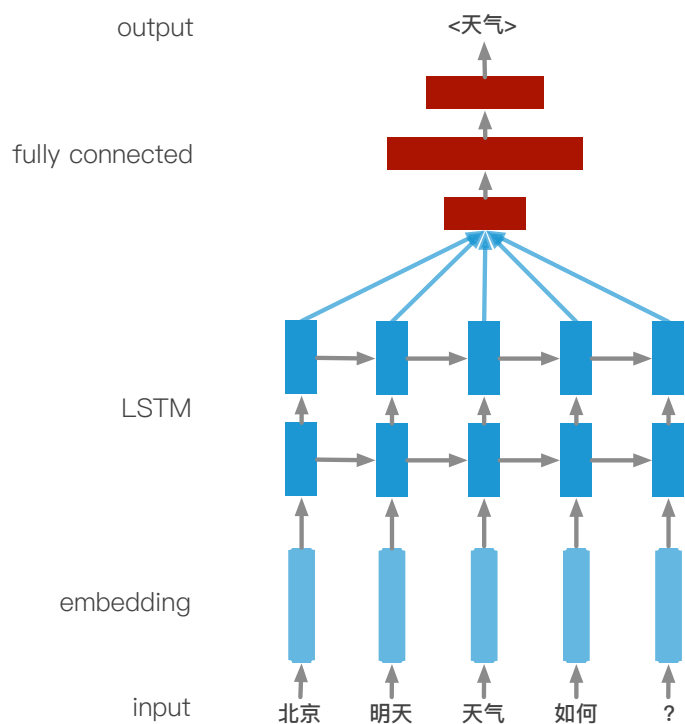
- 亦称为**槽位提取**
 - 序列标注问题
 - HMM/CRF
 - LSTM
 - BiLSTM-CRF



Route-Bot——领域/意图检测

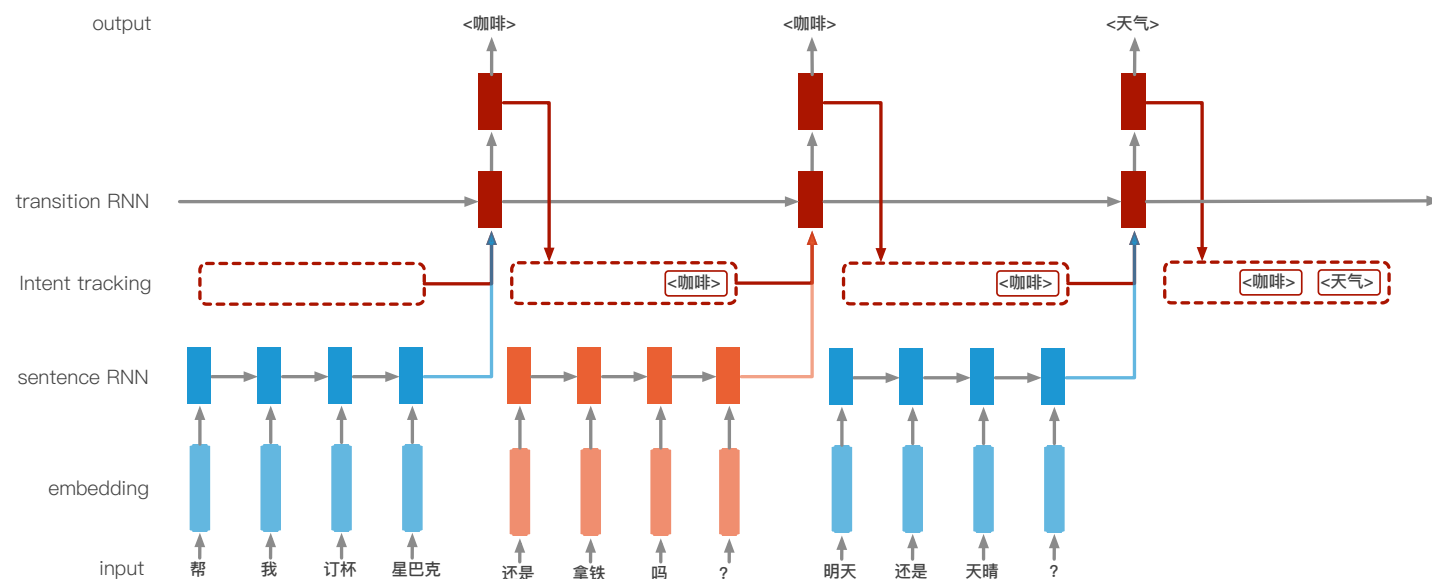
- 单轮识别模型

- RNN/CNN + DNN

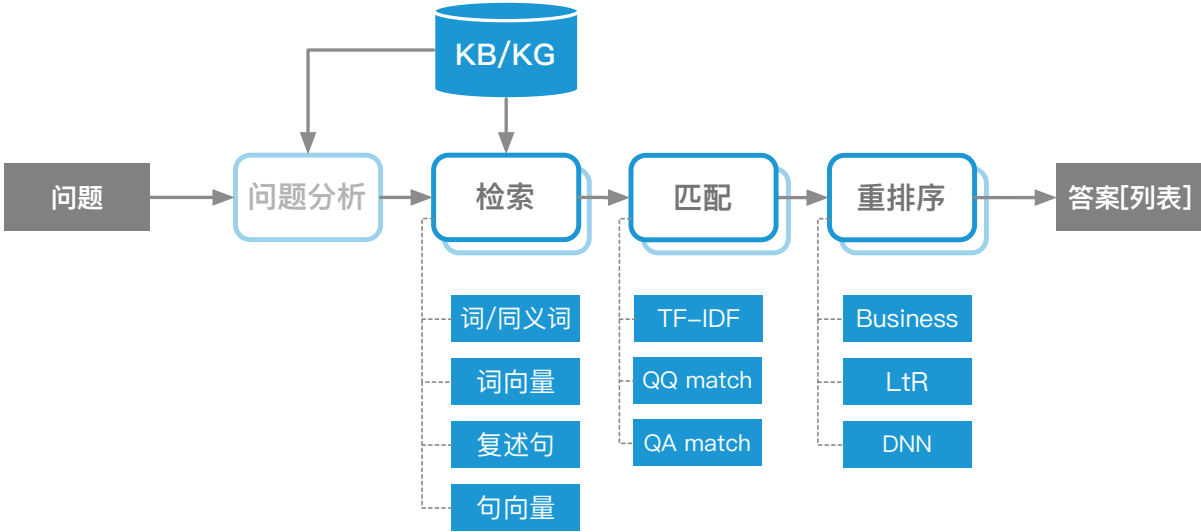


- 多轮识别模型：多轮对话

- RNN/CNN + RNN



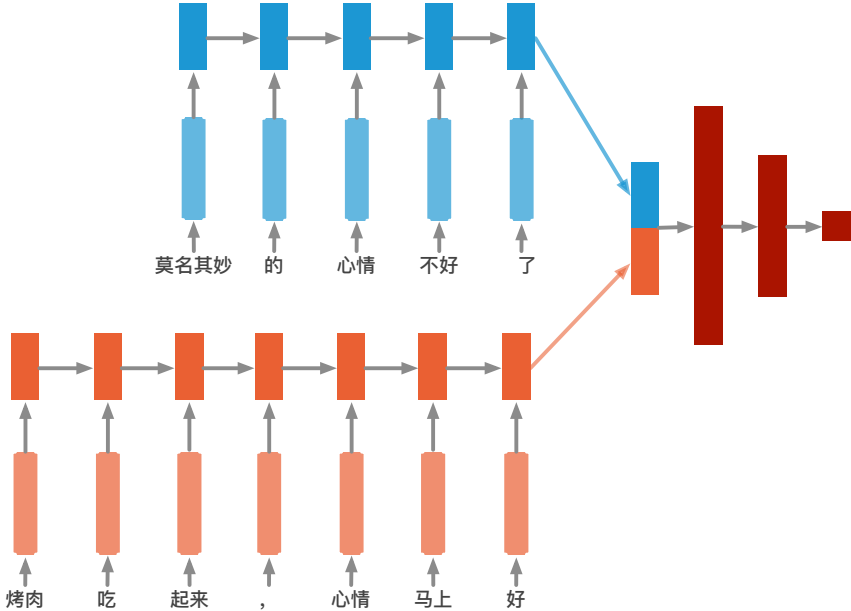
FAQ-Bot



- 特色

- 语义匹配
- 问题带背景
- 答复带参数

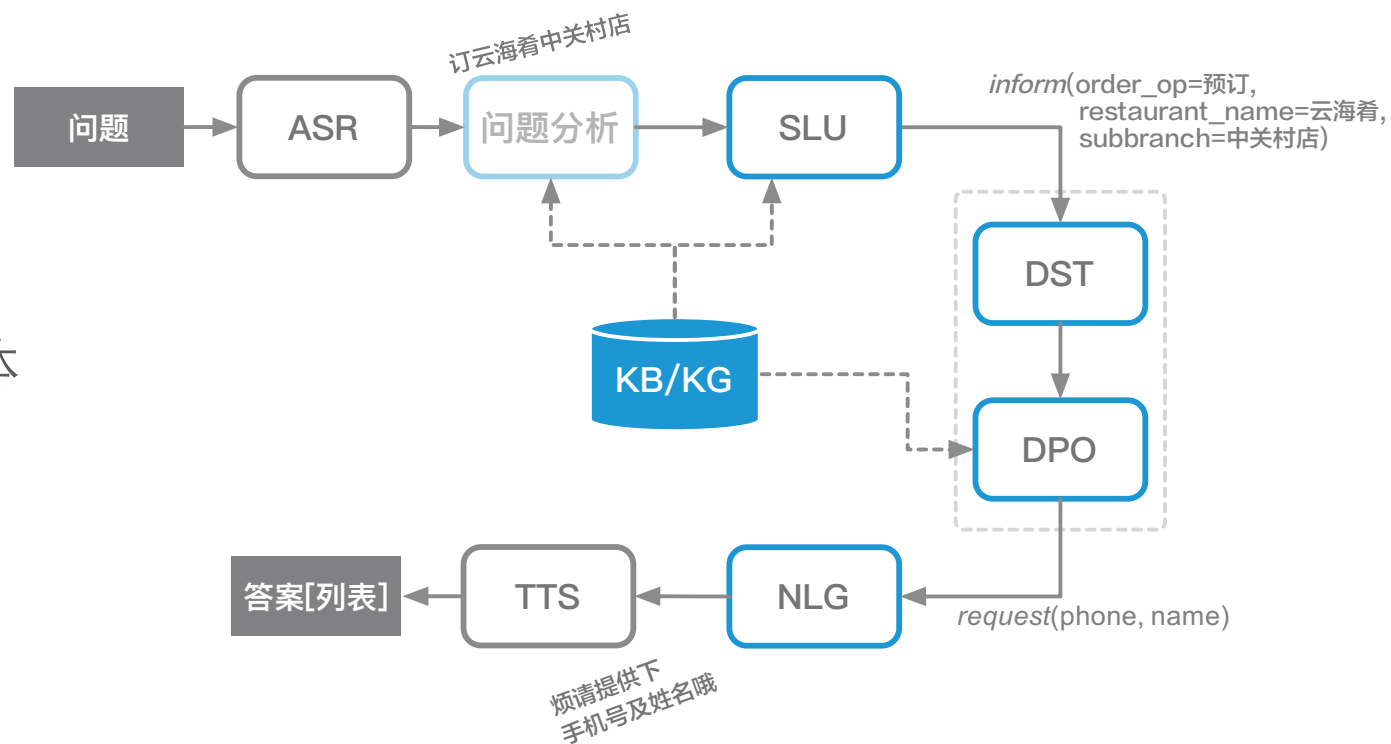
- 匹配模型
 - CNN/RNN + DNN



Task-Bot

- 系统框架

- SLU -> DST -> DPO -> NLG

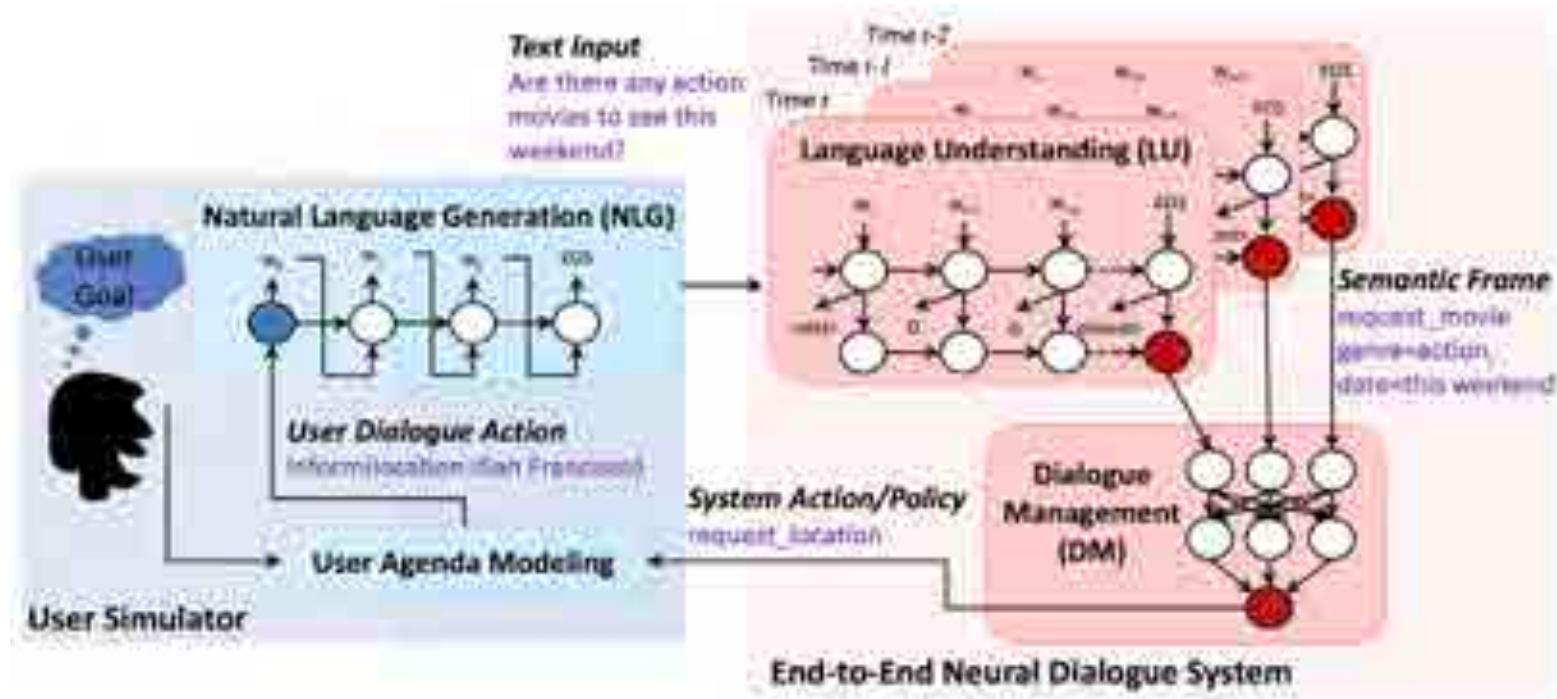


- 特色

- 任务包含20+种实体
- 实际项目中包含实体最多的任务

Task-Bot

- DM: DST + DPO \rightarrow DRL
 - <https://github.com/MiuLab/TC-Bot>



End-to-End Task-Completion Neural Dialogue Systems

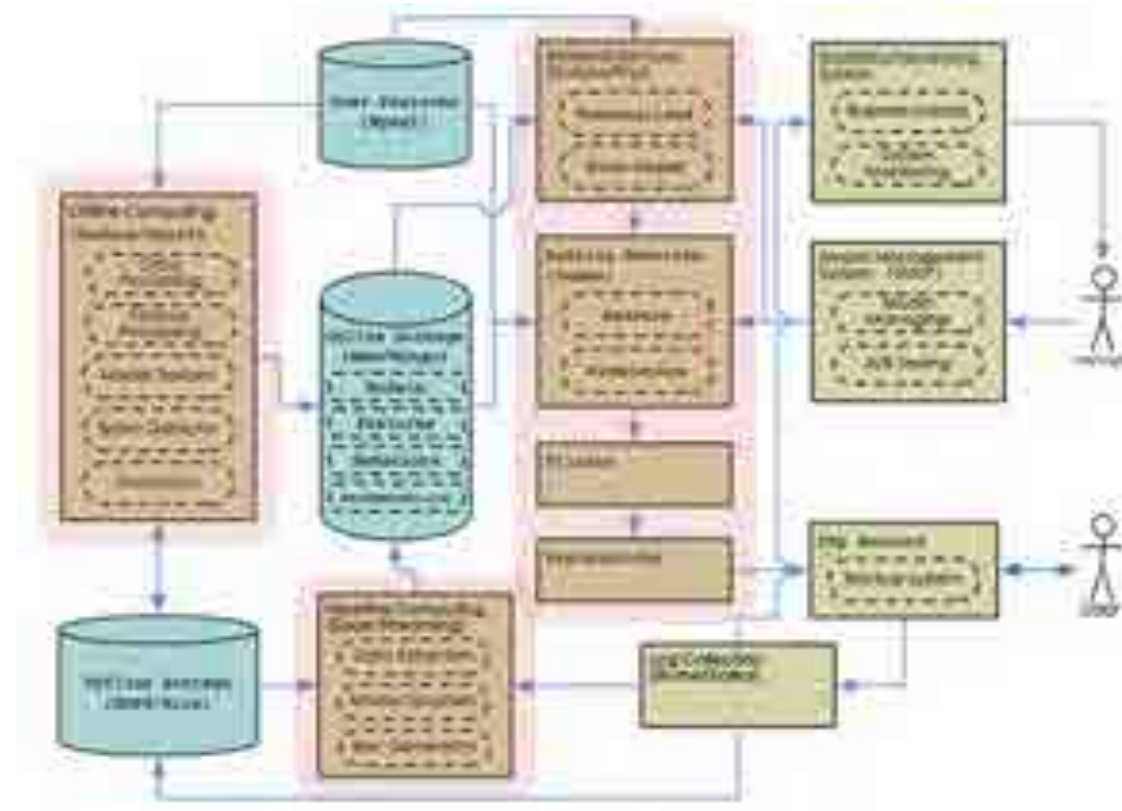
Rec-Bot

- 帮助三方共赢



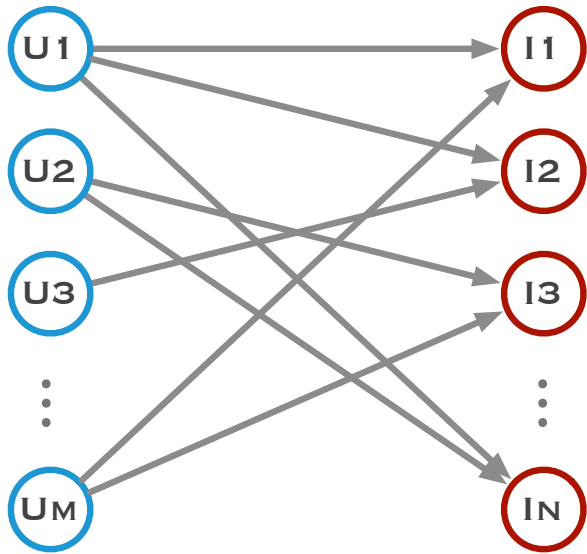
Rec-Bot

- 系统框架：线下 + 近实时 + 实时

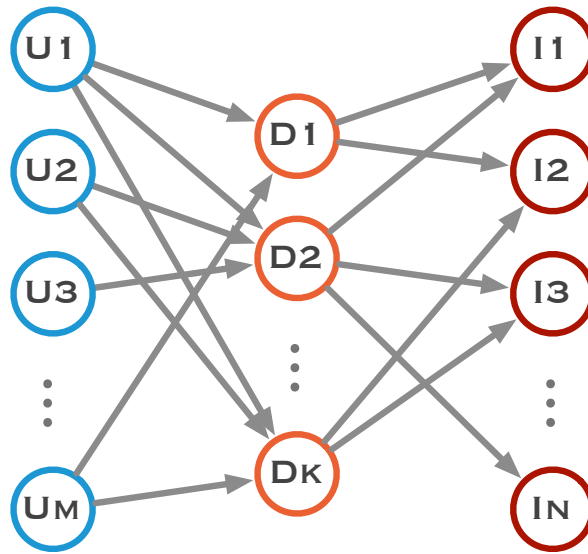


Rec-Bot

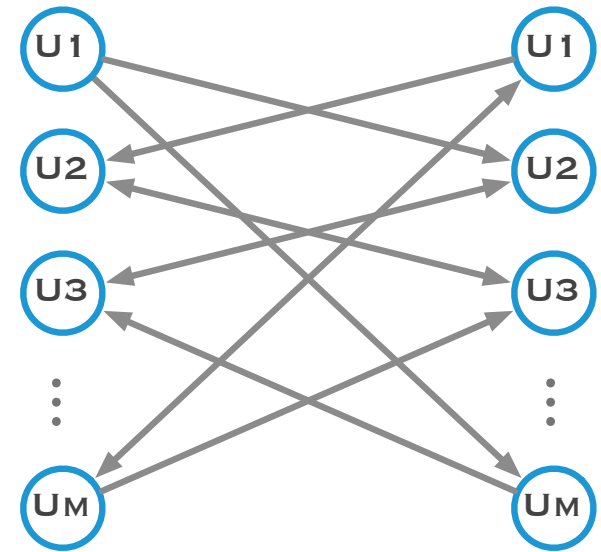
- 各种交互数据



用户-产品



用户-需求-产品



用户-用户

Rec-Bot

- 代表性模型

- Item/User-based kNN

- MF

$$F(\mathbf{P}, \mathbf{Q}) = \frac{1}{2} \sum_{(u,m) \in \mathcal{P}} \left\{ (r_{u,m} - \mathbf{p}_u^T \mathbf{q}_m)^2 + \lambda (\|\mathbf{p}_u\|_2^2 + \|\mathbf{q}_m\|_2^2) \right\}$$

- FM

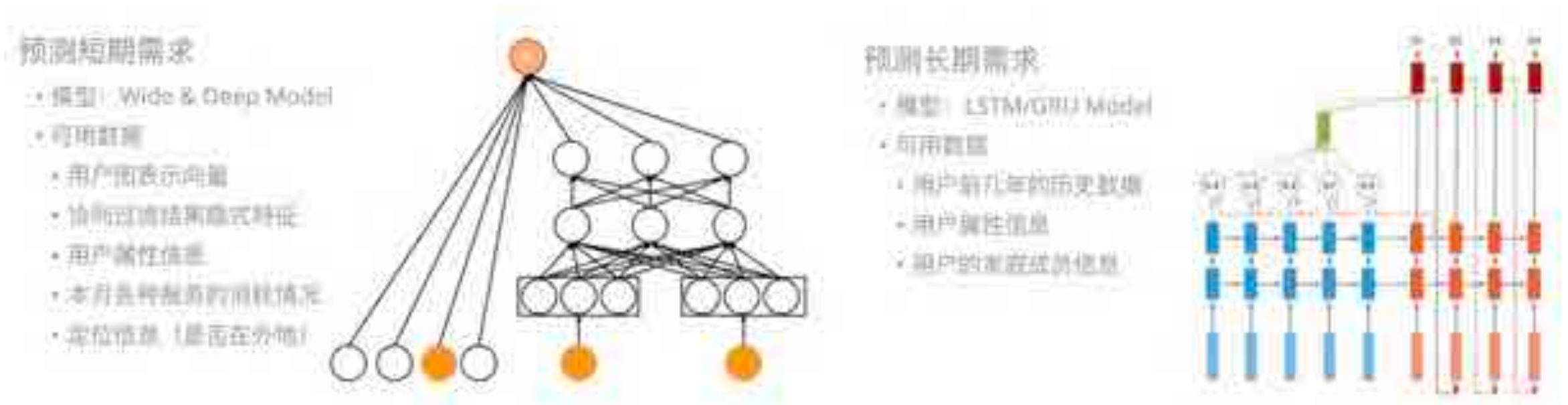
$$\hat{y}(\mathbf{x}) := w_0 + \sum_{i=1}^p w_i x_i + \sum_{i=1}^p \sum_{j>i}^p \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j$$

- DNN/CNN/RNN

Rec-Bot

- 表征用户

- 利用表示学习获得用户相关的各种特征，再把这些表示结果用于预测用户的短期需求和长期需求



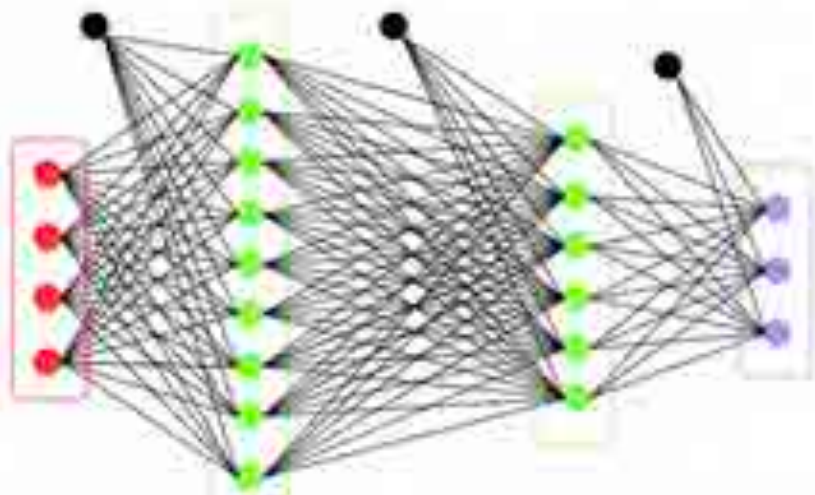
Rec-Bot

- 表征产品

- 每个产品包含的服务以及价格等基本信息，可以显性刻画此产品；而用户对产品的购买等行为数据，可以隐性地刻画此产品

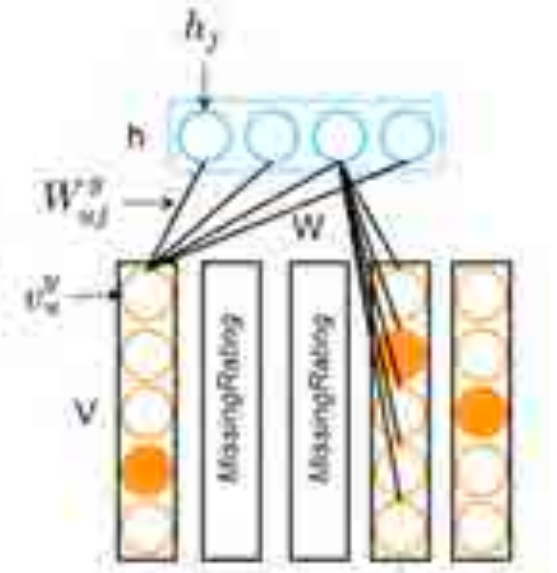
产品的显性刻画

- 模型: DNN Model
- 可用数据
 - 产品使用周期
 - 触发的推荐内容
 - 价格等基础信息



产品的隐性刻画

- 模型: RBM Model
- 可用数据
 - 历史上用户对产品的购买
 - 消费数据

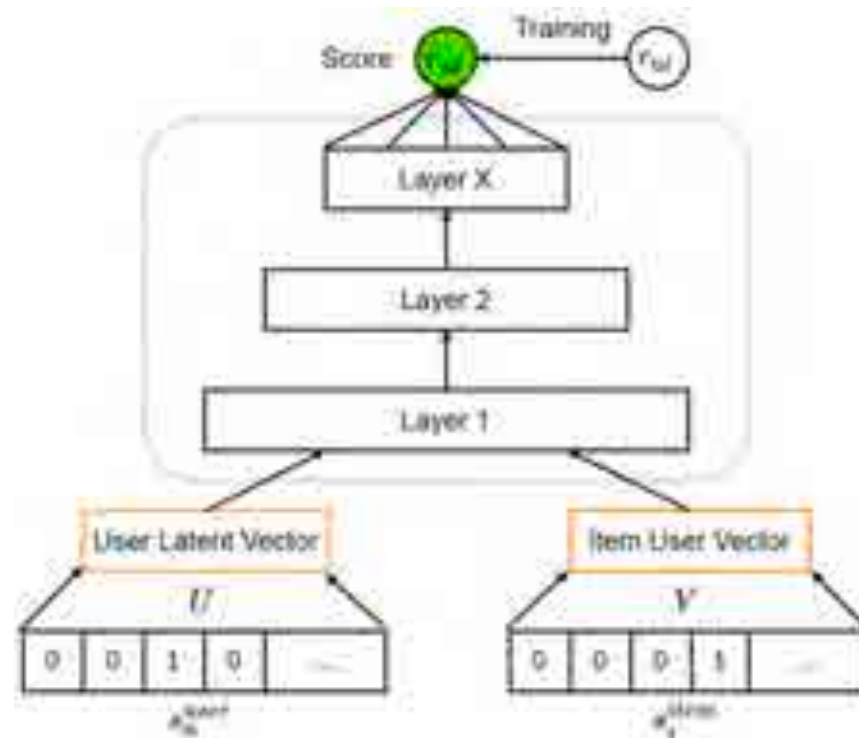


Rec-Bot

- 表征用户×产品
 - 在获得了用户与产品的表征向量，以及用户的短期与长期需求后，使用推荐模型从历史数据中学习模型

- 目标函数借鉴Learning to Rank

- Point-wise
- Pair-wise
- List-wise



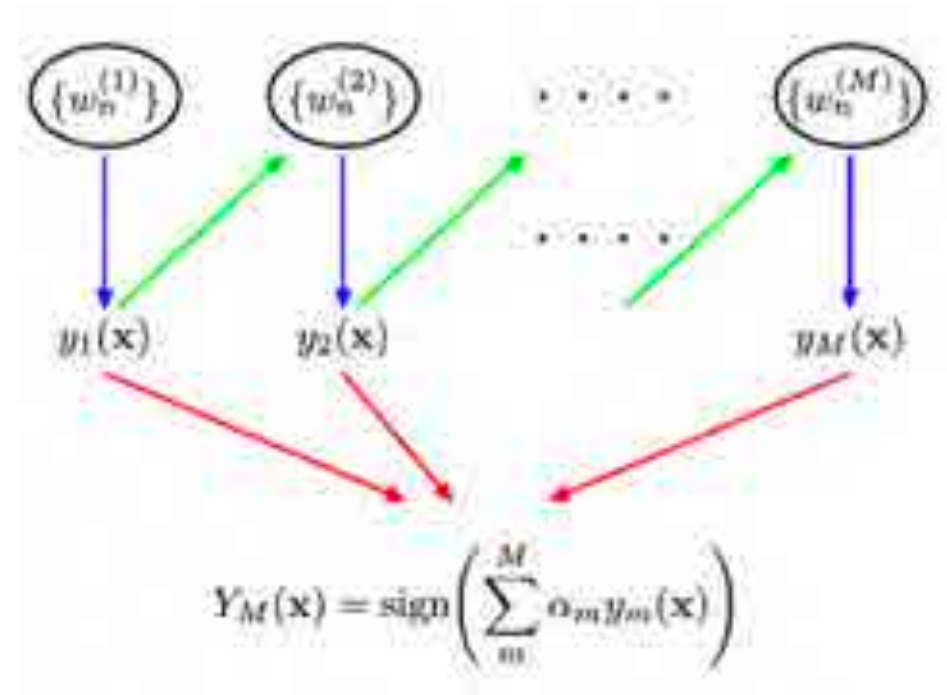
Rec-Bot

- **融合 (Ensemble)**

- 兴起于Netflix Prize竞赛，已成业界标配

- 常用组合方法

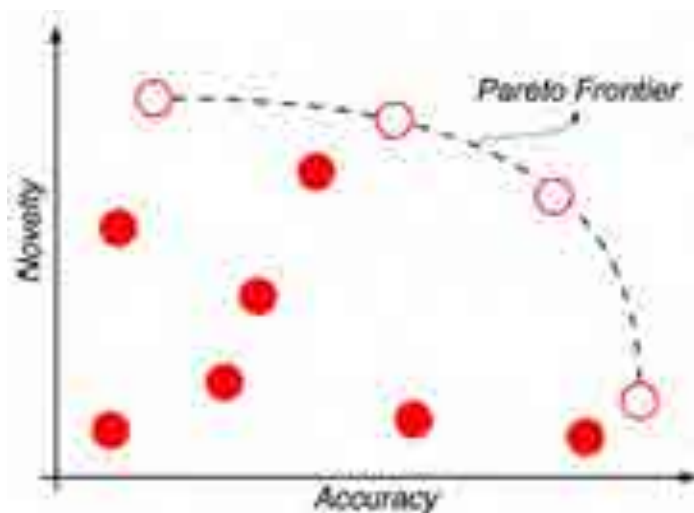
- **预测值组合**：组合多个算法的预测值
 - LinearReg, NN, GBDT, RF
- **特征扩充**：一个算法的输出作为另一个算法的输入特征
- **切换**：不同算法间相互切换
 - DT, Rule-based



Rec-Bot

- 准确度之外的其他指标
 - 覆盖面 (Coverage) : 被推荐的物品占所有物品的比例
 - 多样性 (Diversity) : 覆盖用户各方面的兴趣, 而不仅是一部分
 - 新颖性 (Novelty) : 推荐长尾 (小众) 部分的物品
 - 惊奇性 (Serendipity) : 用户不期望系统能推荐, 但用户真感兴趣
- 多目标
 - Pareto边界: 在保证其他目标在一定约束范围内, 优化主要的一个目标
 - 多个目标综合为单个目标

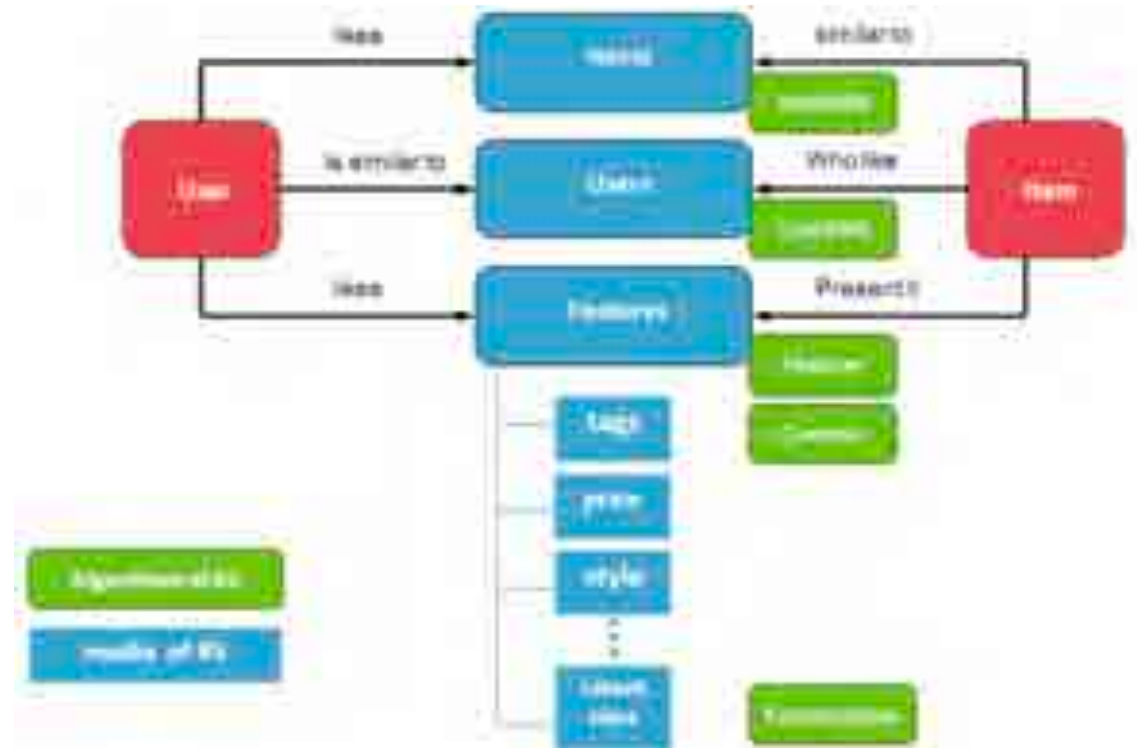
- 技术指标 \neq 业务指标



Rec-Bot

- 推荐理由

- 使用推荐服务时，用户往往没有明确需求，推荐系统需要提供能说服用户的理由
 - 通过产品提供的服务
 - 通过服务消耗可以推断用户兴趣，进而产生推荐理由
 - 通过优惠信息
 - 通过具体的节省计算



Chitchat-Bot

- 检索 + 生成式

- 检索

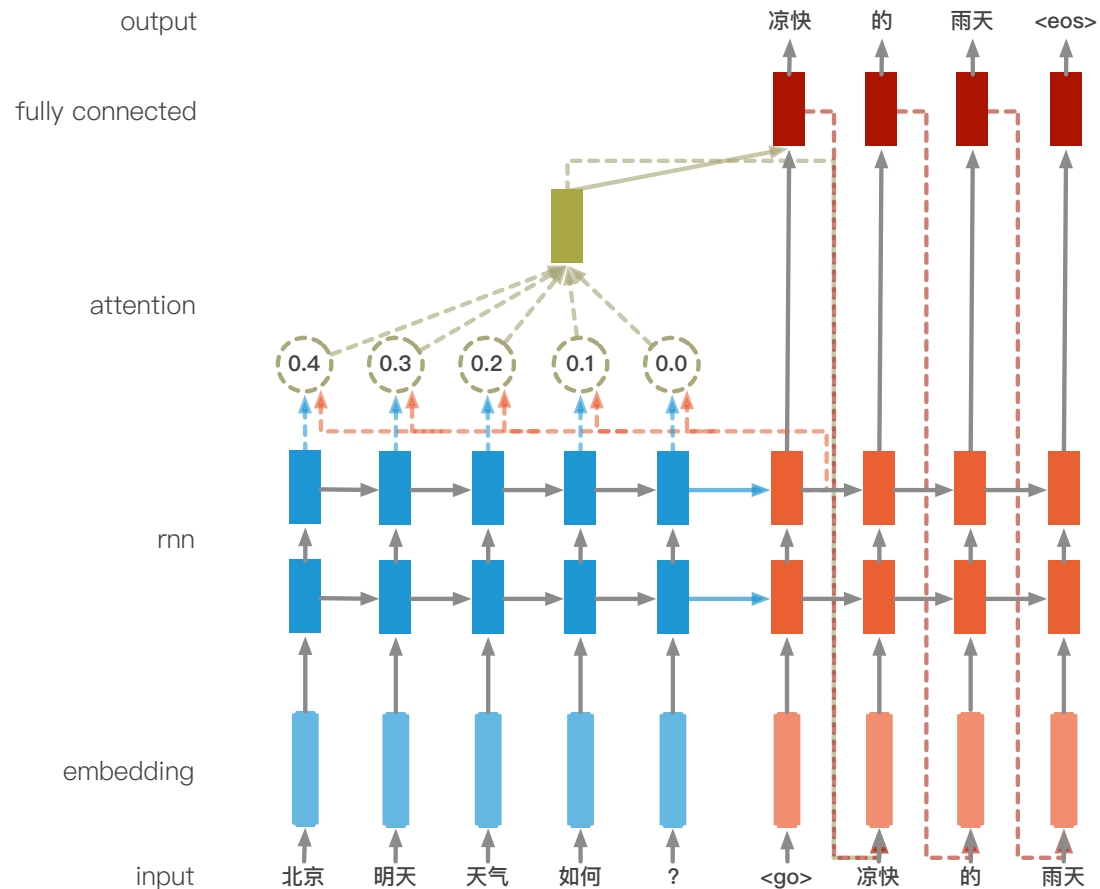
- 保证可靠性：允许客户添加qa对，秒级生效

- 生成式

- 保证100%召回：可以回答任何问题，降低知识库维护门槛

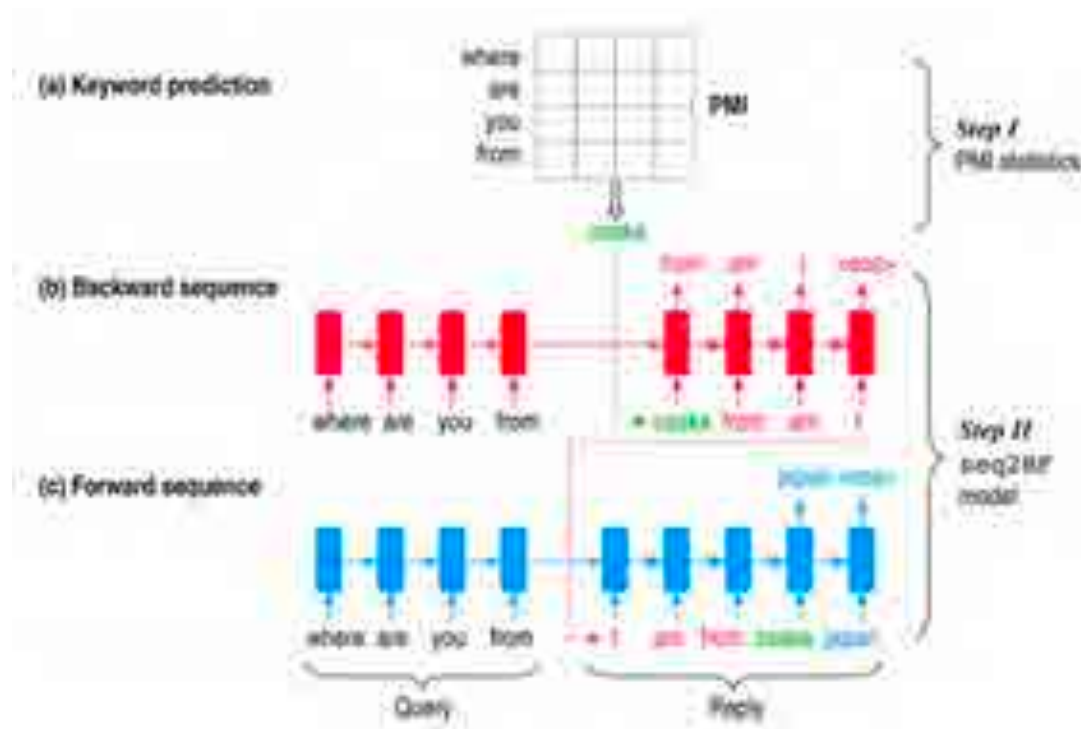
- 特色

- 使用coverage机制保证生成式语句不重复
 - 使用copy机制保证稀有词的产生
 - 使用MMI提升生成结果的多样性



改进seq2seq

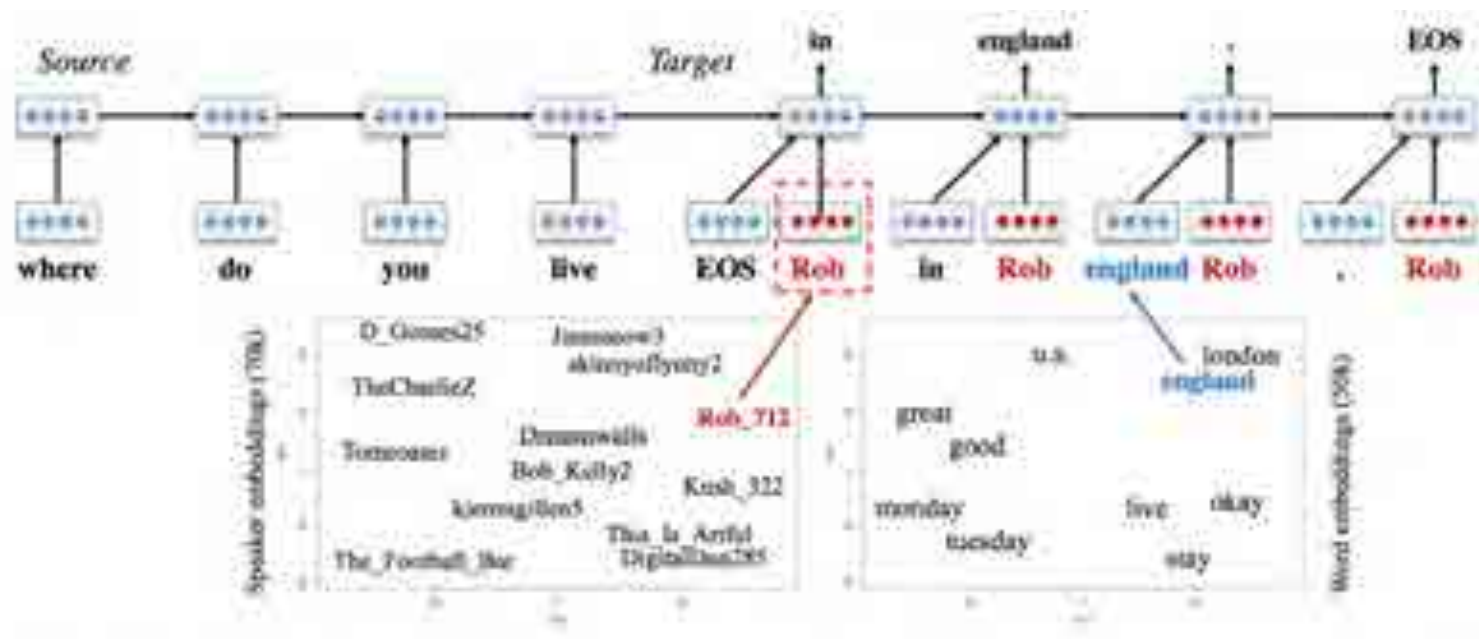
- 注入额外的信息
 - 背景信息：主题关键词



sequence to backward and forward sequences

改进seq2seq

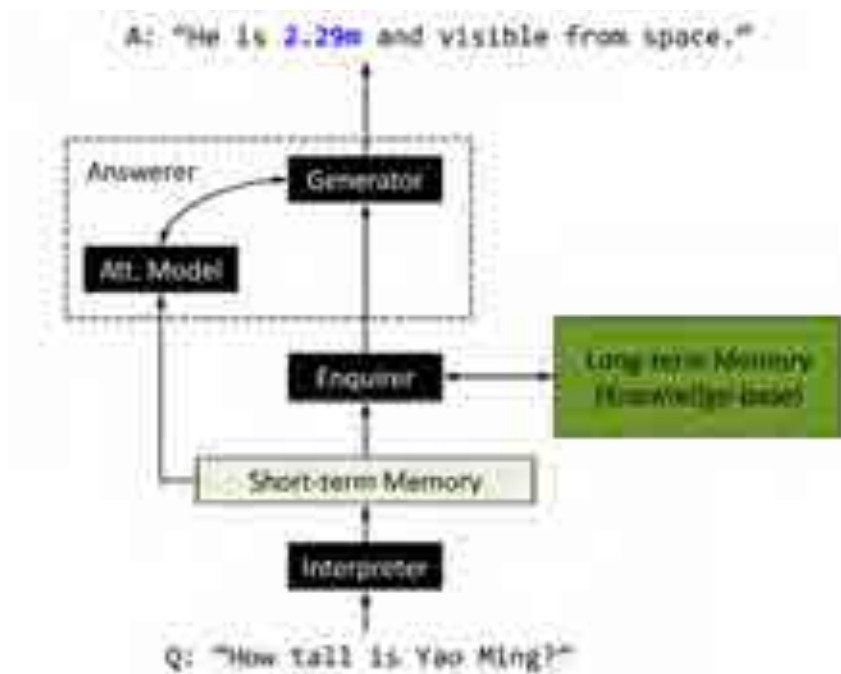
- 注入额外的信息
 - 对话双方的个人信息



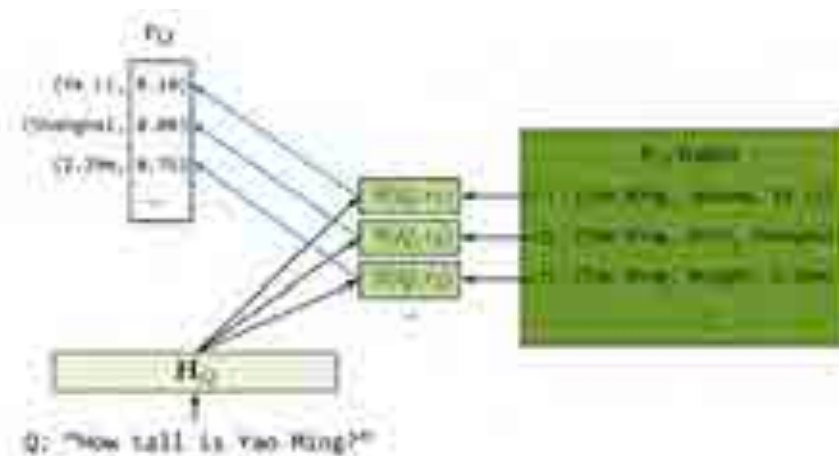
A Persona-Based Neural Conversation Model

改进seq2seq

- 注入额外的信息
 - 常识：知识库/图谱

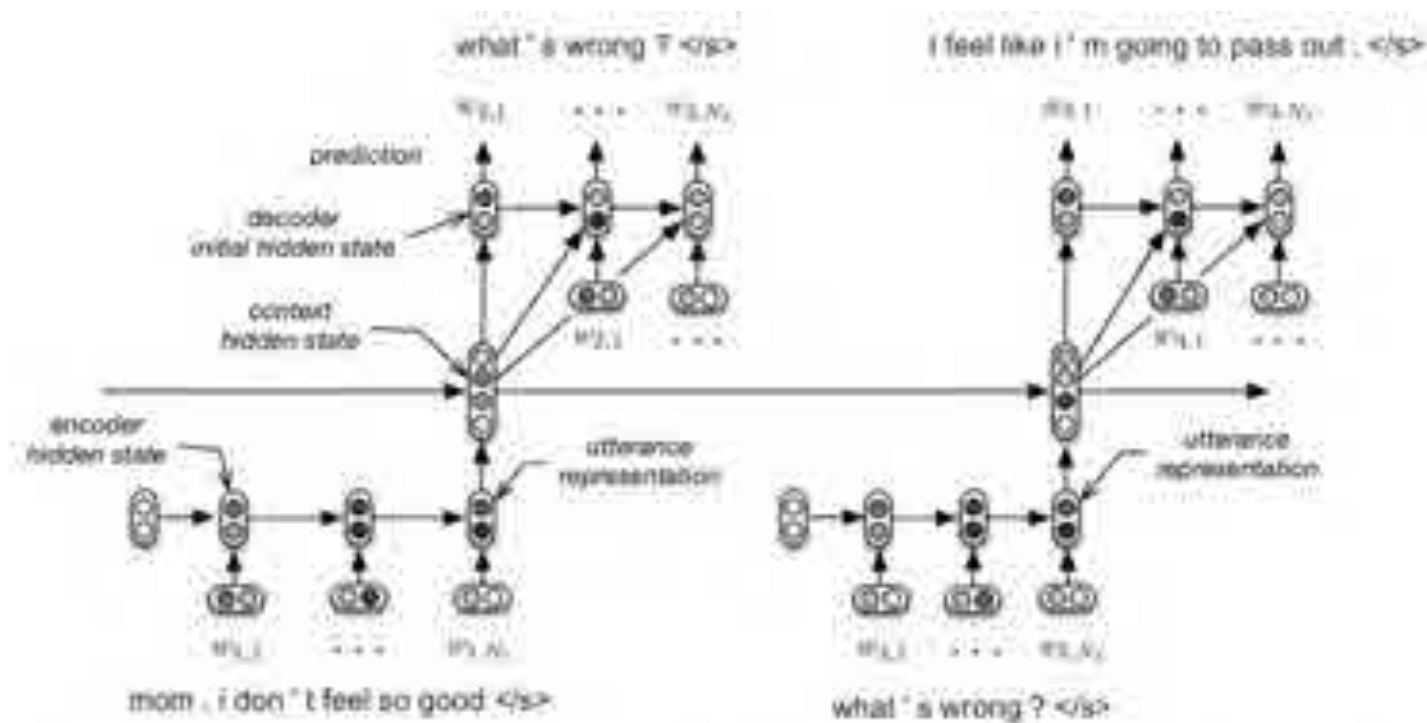


Neural Generative Question Answering



改进seq2seq

- 注入额外的信息
 - 上下文



Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models

CUI / GUI

商业实践

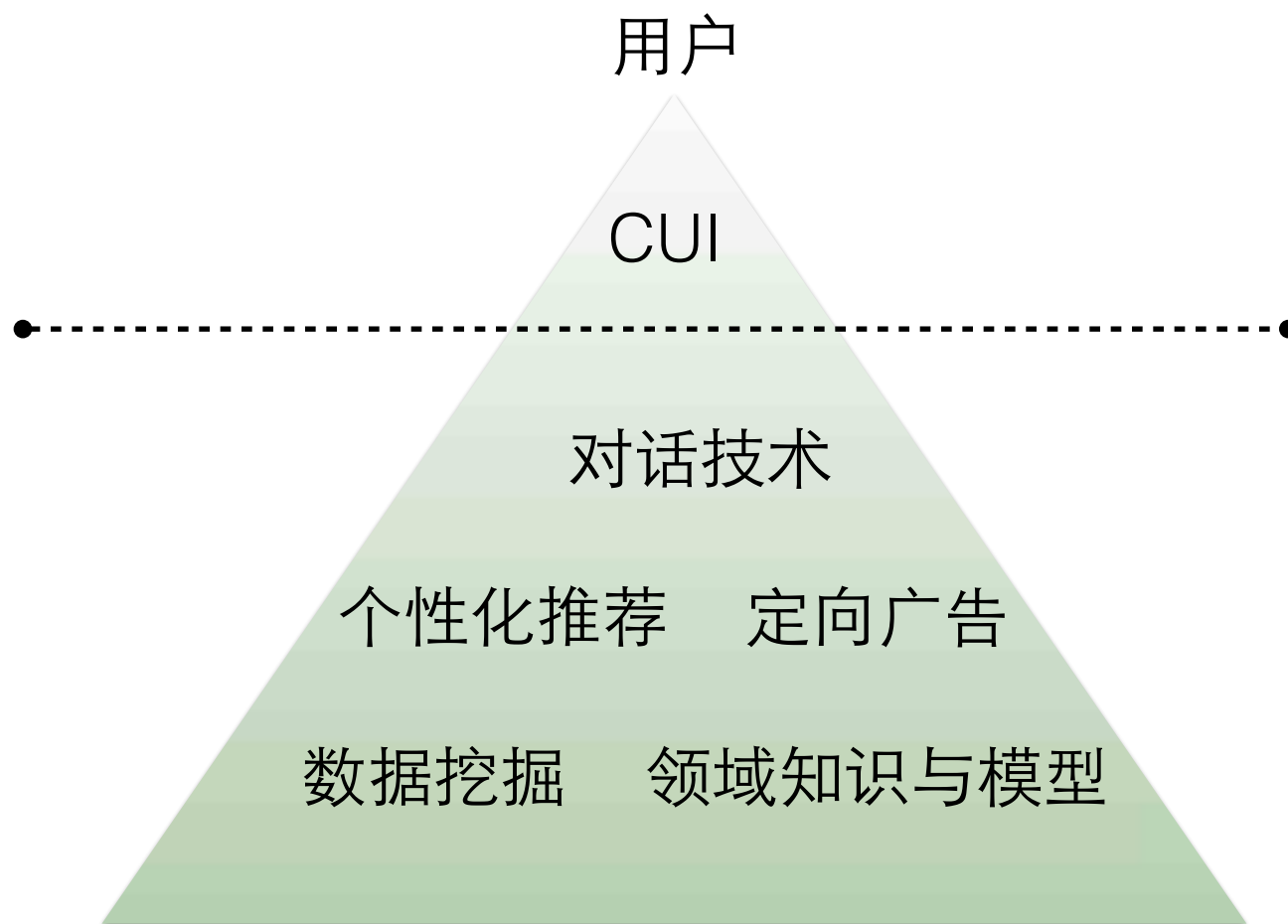
CUI / GUI

效率： 信息展示的广度和深度

感受： 时间感与空间感对沟通的影响

预期： 个性化、参与感、自我学习与进化

对话作为粘合剂



对话的商业原则

用户价值： 通过对话高效解决问题

稀缺性： 成为细分领域的关键业务环节

技术成熟度： 选择合适的技术

对话商用场景

- 清晰的知识结构和边界
- 非标准化服务，信息不对称
- 能够通过数据积累提升服务质量
- 能够建立知识和技术壁垒

几个有趣的场景

智能投顾、保险、理财等销售转化

对话式的发现、解释和推荐

消息平台的 Landing Page

总结

趋势是大的

技术是不够的

商业是要探索的



吴金龙 3

2017 03 23



微博: [@breezedeus](#)

博客: [breezedeus.github.io](#)

Thanks!

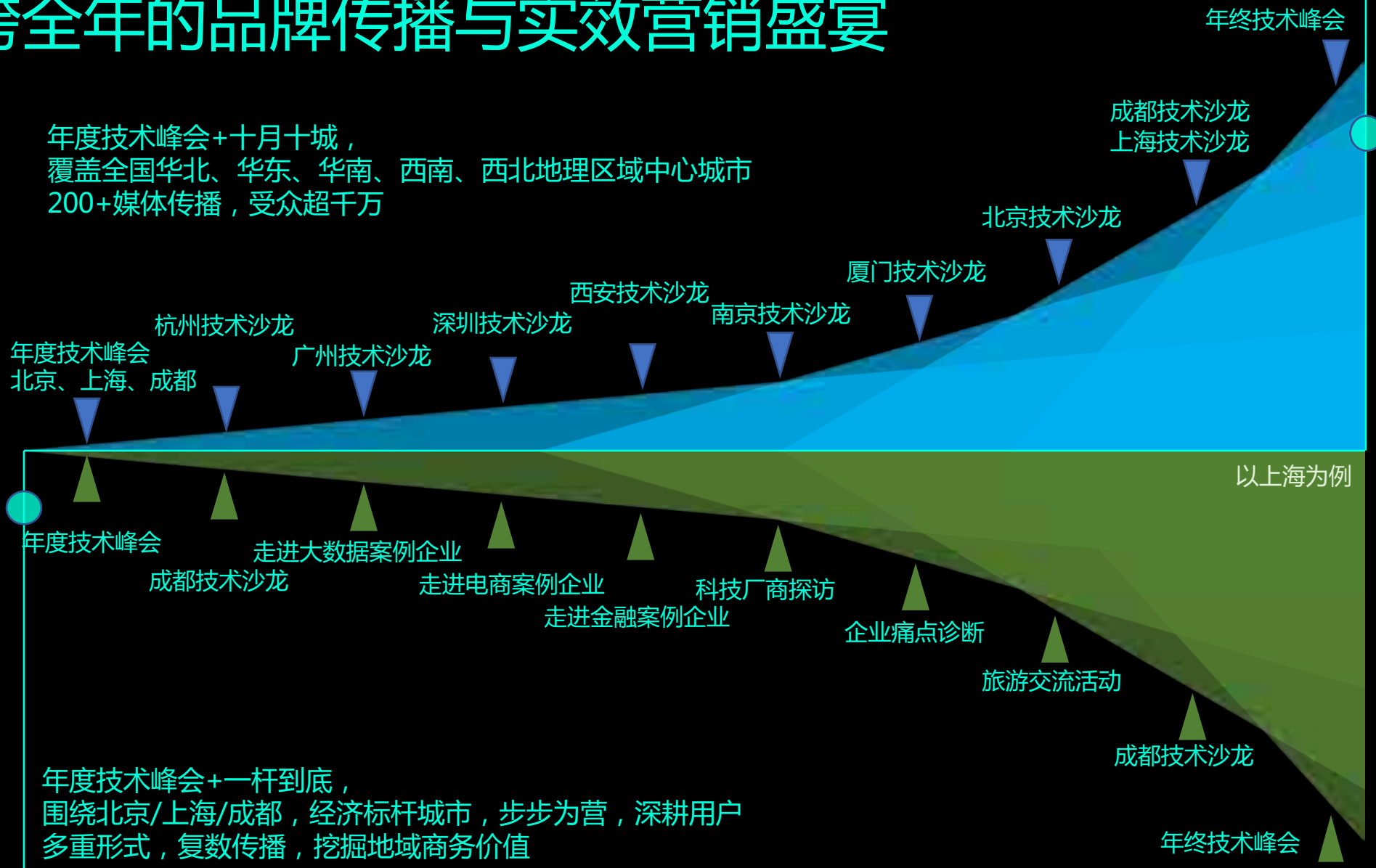
Your business is in good hands.



横跨全年的品牌传播与实效营销盛宴

同城合纵

全域连横



合纵连横，在中国开发者群体中缔造品牌营销奇迹



中生代技术

FRESHMAN TECHNOLOGY



ArchData技术峰会全国巡回
上海9月, 北京9月, 成都10月, 南京10月,
长沙11月, 广州11月
中生代咨询内训
技术架构, 研发管理, 敏捷开发, 大数据
微服务, AI, 机器学习
中生代人才内推
对接研发主管, 内推精准人才