

智变中的美团客服

刘学梁 20180916



基础研发平台 | AI平台NLP中心

目录

01 客服系统简介

演变中的客服系统，对话系统，智能客服机器人，美团业务简介，美团客服系统

02 智变之路

智变基础，智变途径，系统架构，QABot，TaskBot，ChatBot，自主学习流程

03 落地效果

QABot线上效果，TaskBot线上效果

04 总结

人机协同学习，加速进化

1.1 演变中的客服系统

智能化是客服系统演变的方向，智能客服通过算法、人机协同、自主学习可以极大地提高人效和体验。

话音呼叫中心

纯人工服务模式
只支持语音电话
人效低，成本高

网页在线客服

基于网页会话
支持文本和语音
访客轨迹跟踪
流量统计分析
客户关系管理

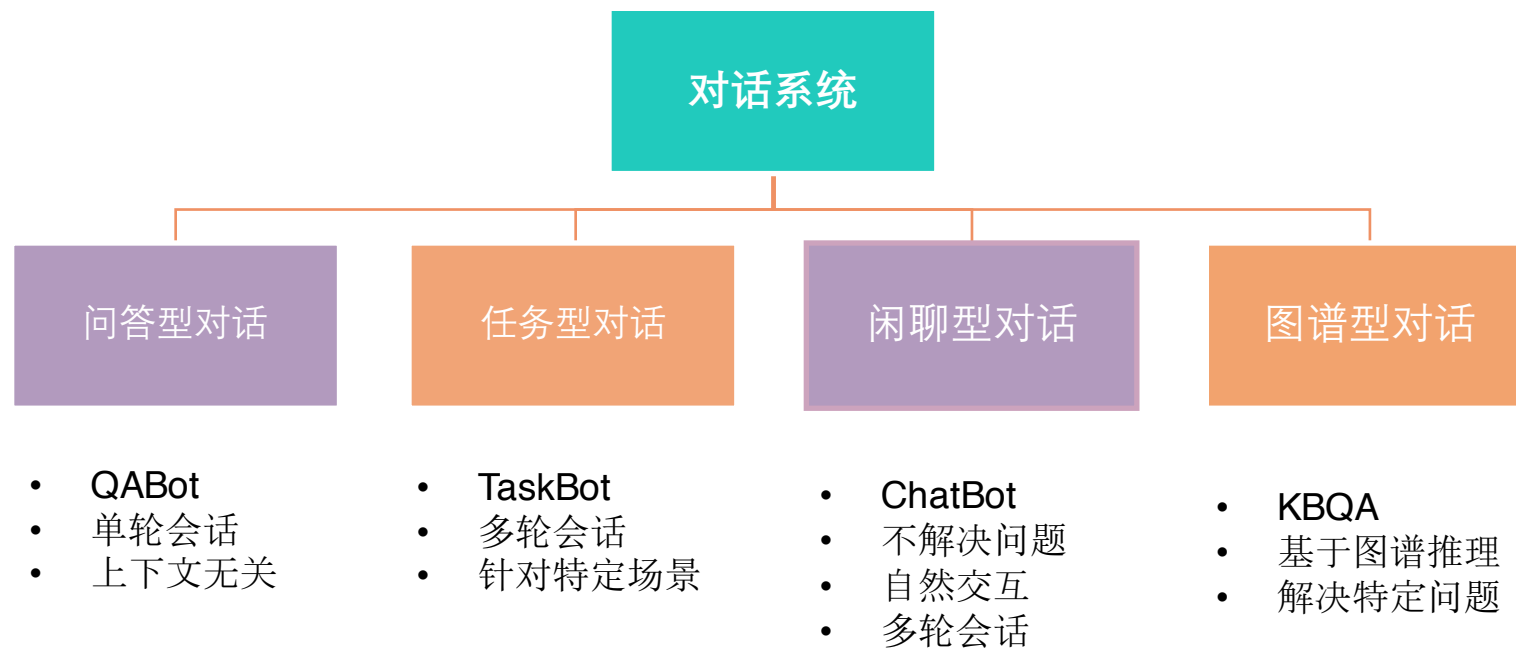
SaaS客服

多渠道接入
丰富的辅助功能
简单知识库管理

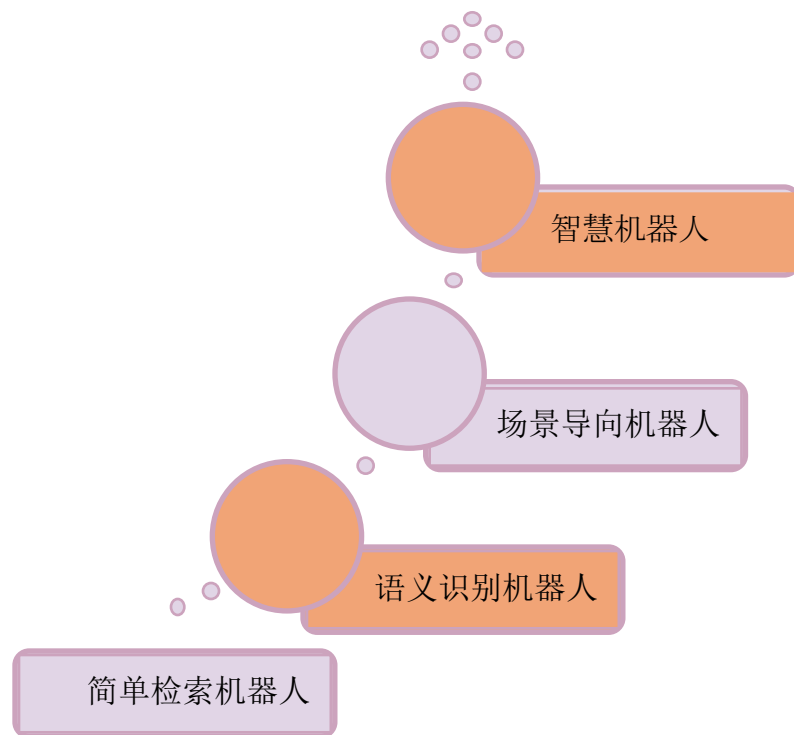
智能客服

自动回复
自助解决问题
人机协同
自主学习

1.2 对话系统简介



1.3 智能客服机器人



- 将对话系统技术融入在线客服体系
- 综合利用语义分析技术、大数据技术以及深度学习技术
- 由过去单一售后服务变为贯穿售前售中售后全客户服务周期
- 智能客服机器人是下一代客服的核心驱动力

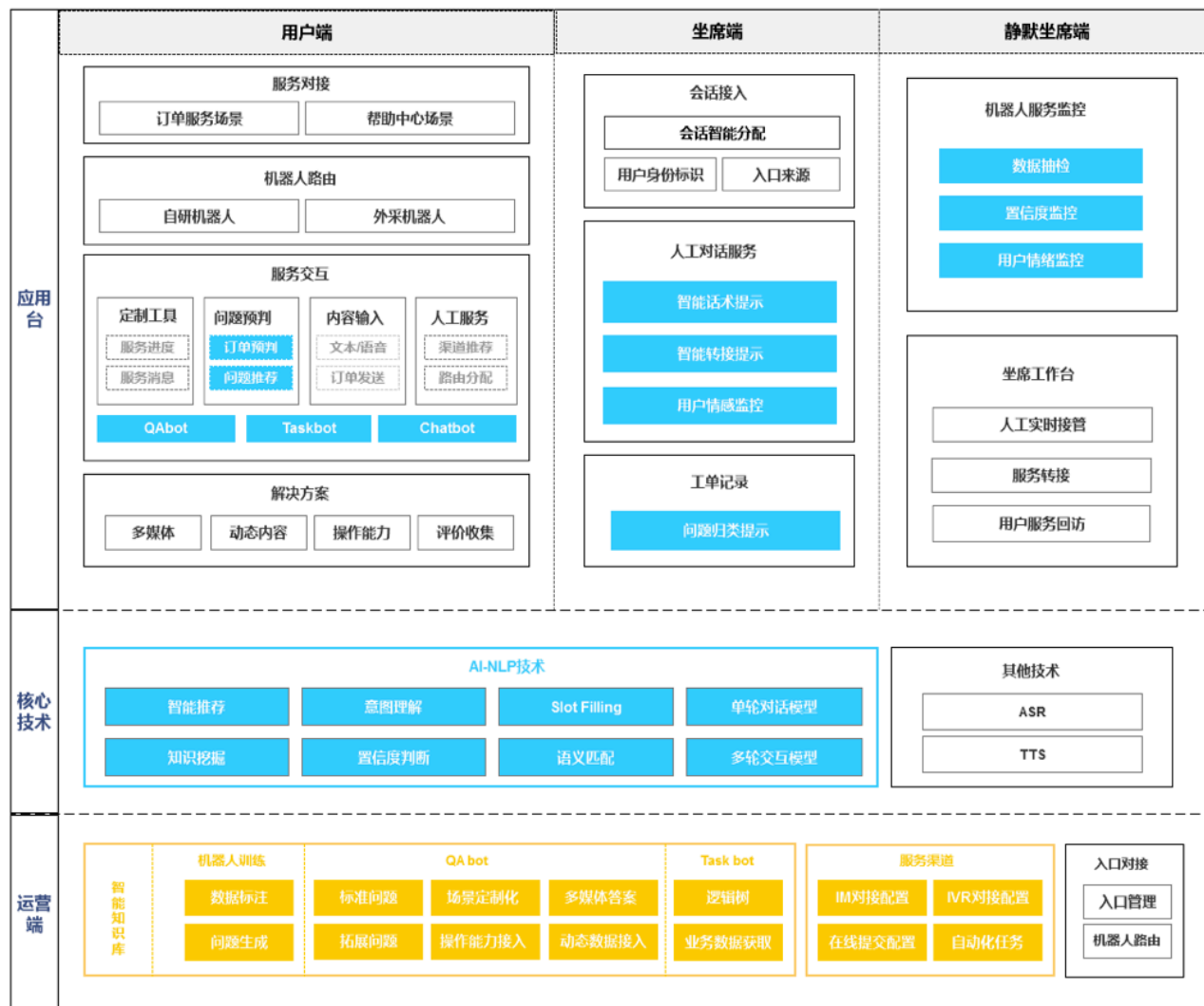
1.3 美团业务简介

- 🍴 美食 >
- 🍷 外卖 >
- 🏨 酒店 **HOT** >
- 🏠 榛果民宿 >
- 🐱 猫眼电影 >
- ✈️ 机票 / 火车票 >
- 🎮 休闲娱乐 / KTV >
- 🏠 生活服务 >
- 💋 丽人 / 美发 / 医学美容 >
- 💍 结婚 / 婚纱摄影 / 婚宴 >
- 👨‍👩‍👧 亲子 / 儿童乐园 / 幼教 >
- 🏃 运动健身 / 健身中心 >
- 🔧 家装 / 建材 / 家居 >
- 📖 学习培训 / 音乐培训 >
- 🏥 医疗健康 / 宠物 / 爱车 >
- 🍷 酒吧 / 密室逃脱 >

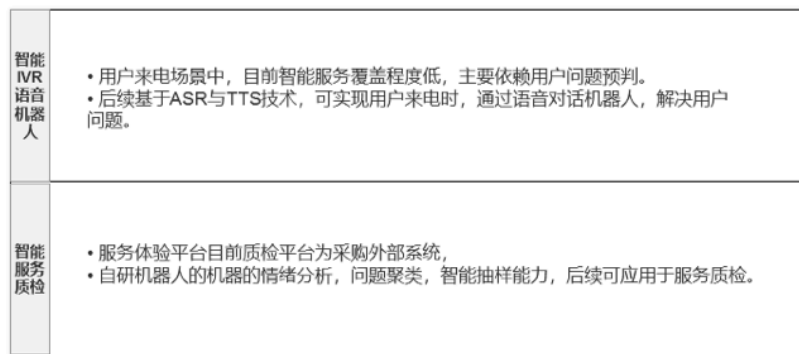
- 业务种类多，吃喝玩乐全都有
- 客户端种类多
- 智能客服是解决有限的客服资源与不断增长的海量用户服务请求矛盾的唯一选择



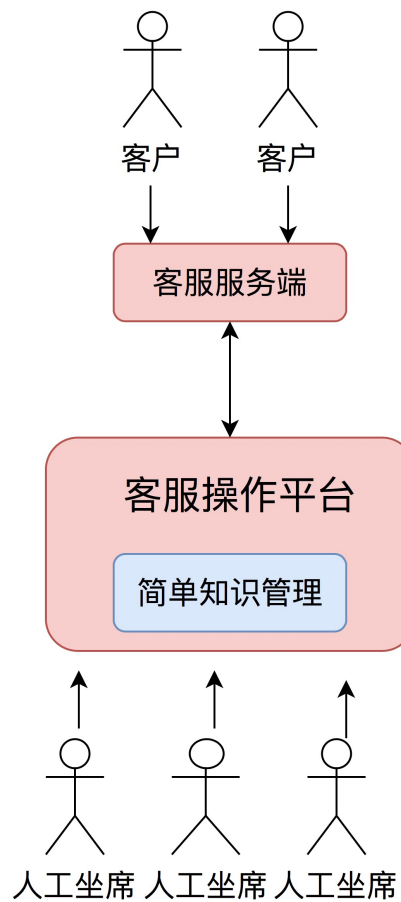
1.4 美团客服系统



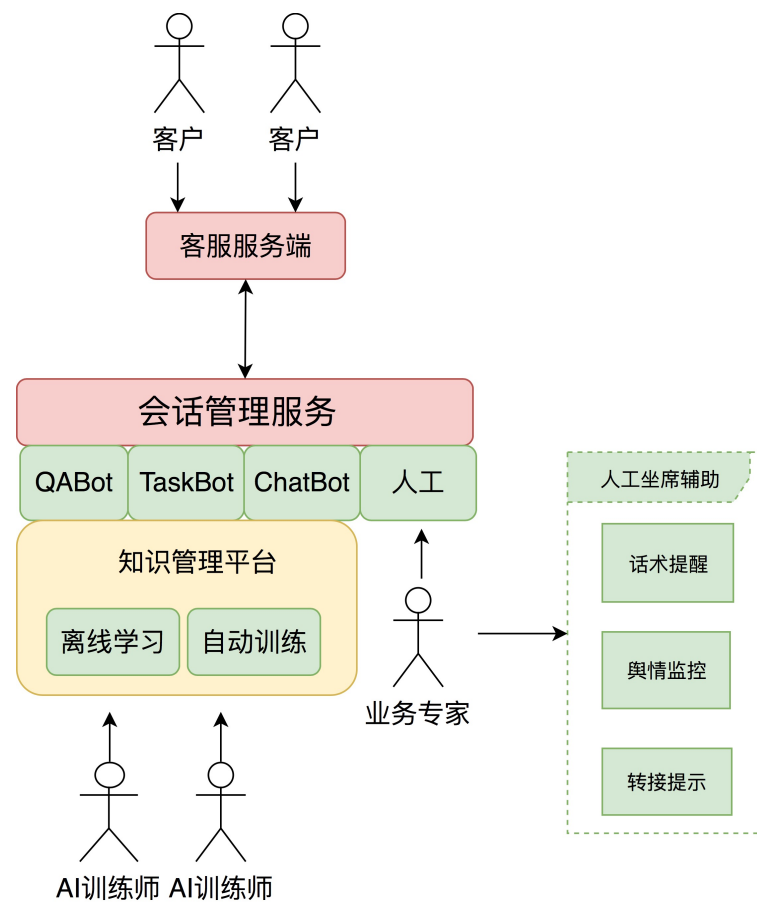
2 应用台后续拓展



2 智变之路

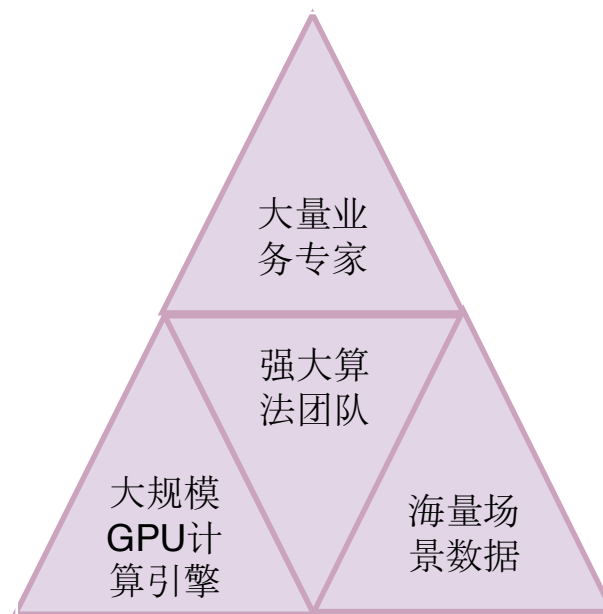
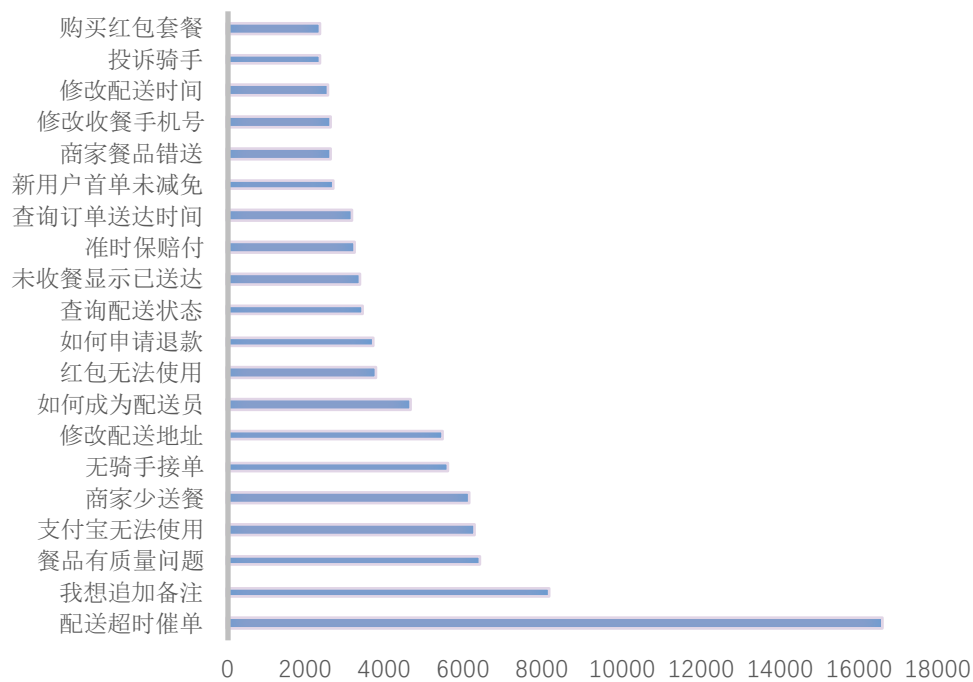


智变



2.1 智变基础

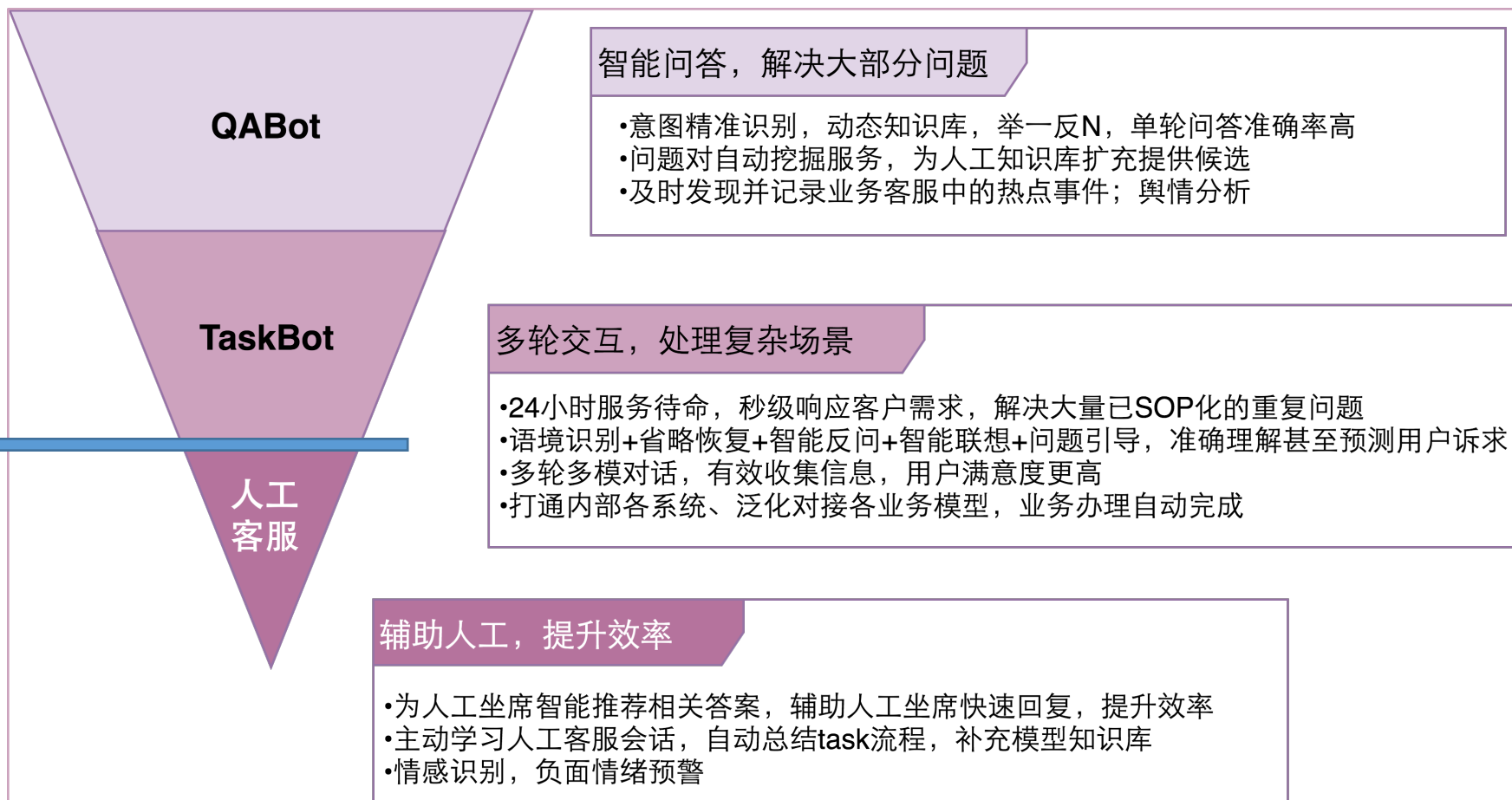
问法数量统计



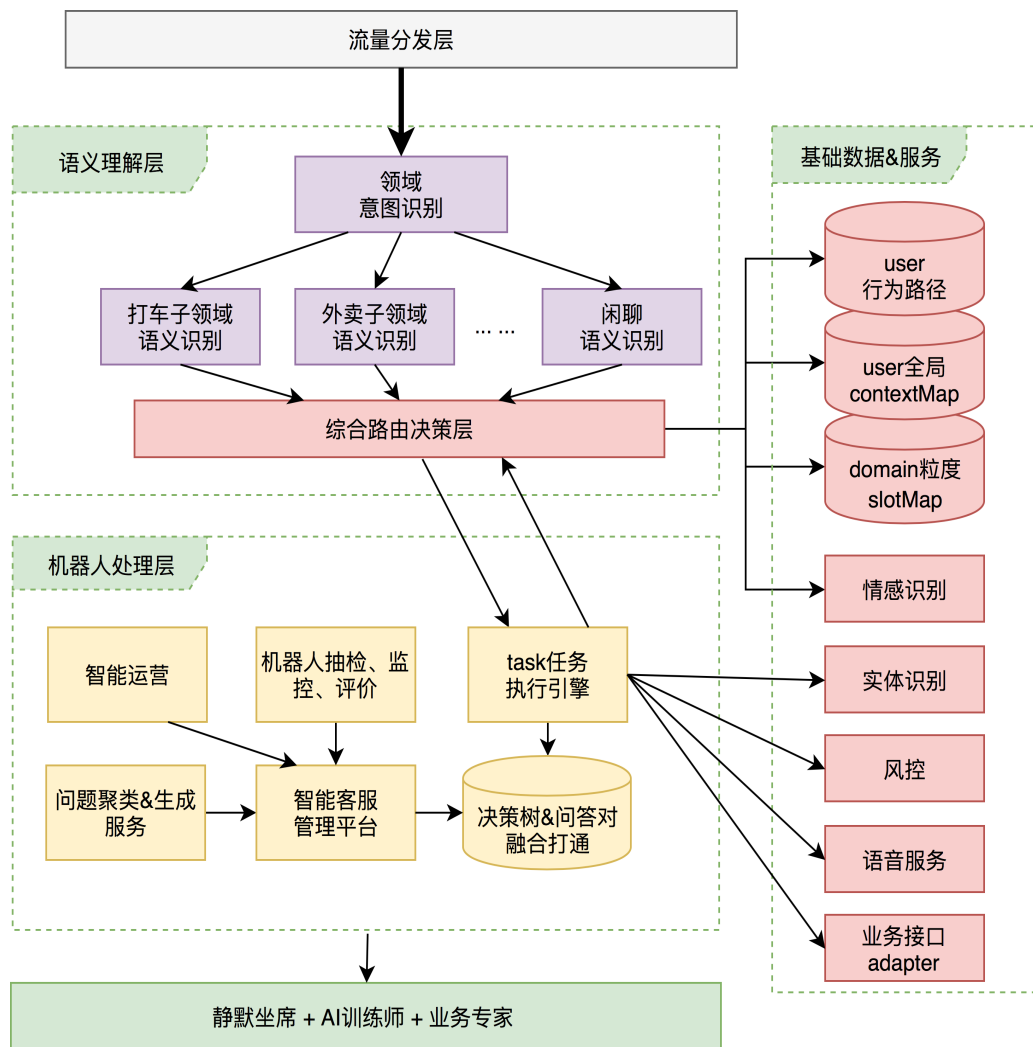
业务数据的特点决定了智变的必要性，团队资源为智变提供了良好的保障

2.2 智变途径

不断向人学习
不断压低这根线



2.3 系统架构



➤ 语义理解层

全生命周期的会话管理。综合用户行为轨迹、会话上下文、用户情绪等信息综合判断用户意图，准确找到问题甚至预测问题。

➤ 机器人处理层

执行单轮会话的答案生成。执行多轮任务会话。主动学习人工客服会话，自动问题归类、自动生成新Task。智能发现不合理Task流程设计，智能优化。

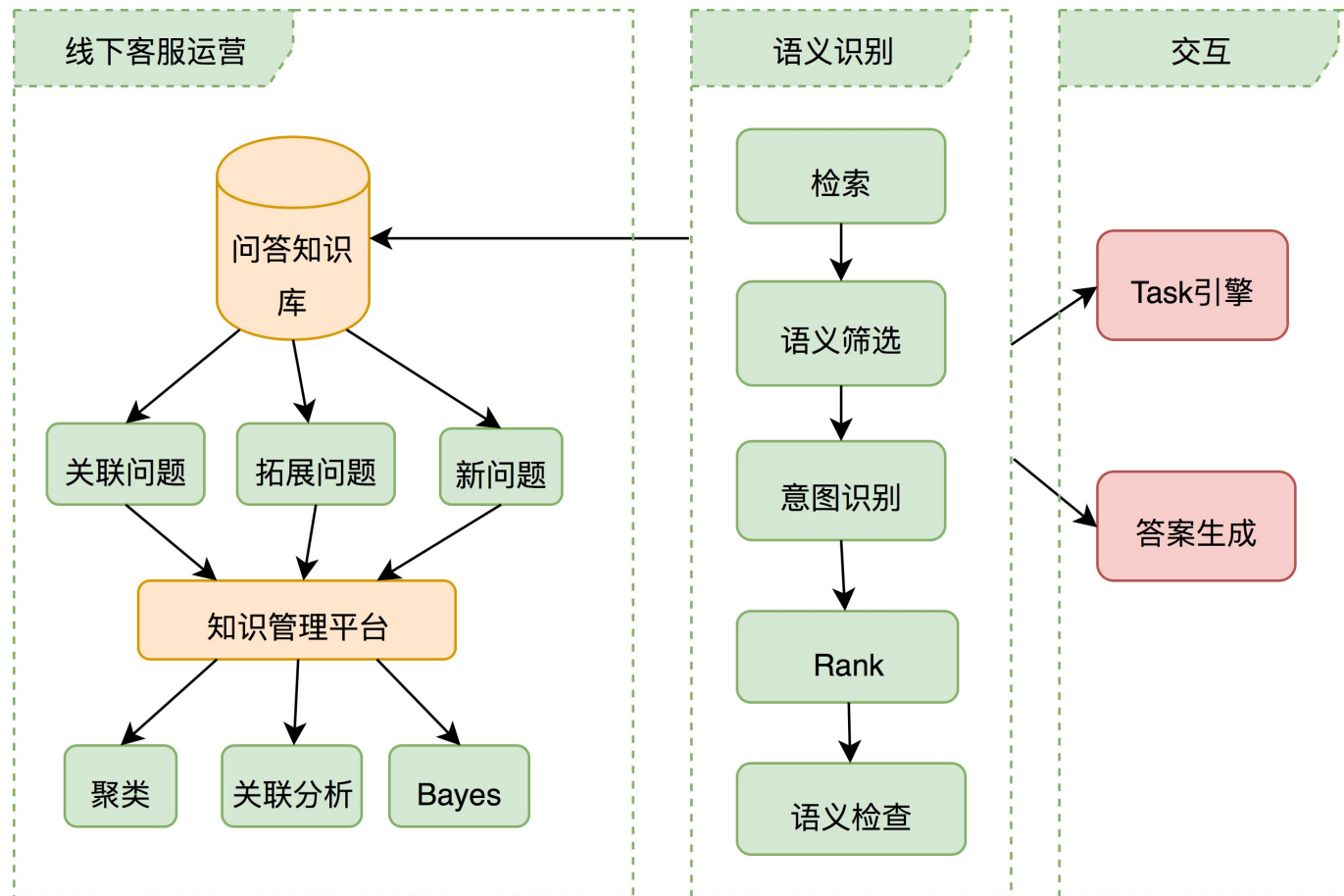
➤ 基础数据&服务层

为会话管理和语义理解提供信息支撑。

➤ 人工层

静默坐席解决机器人解决不了的问题。AI训练师辅助机器人进化。业务专家制定合理的解决方案。

2.4 单轮会话机器人QABOT



- 线上提供语义识别能力
- 线下具有离线学习能力
- 返回精确答案或者推荐问题
- 数据驱动智能
- 少量干预加速进化

2.4.1 语义识别

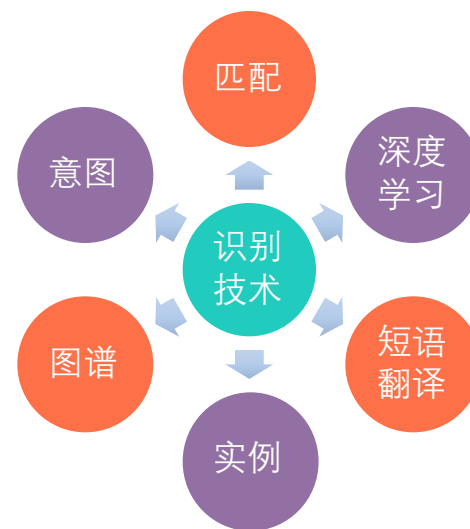
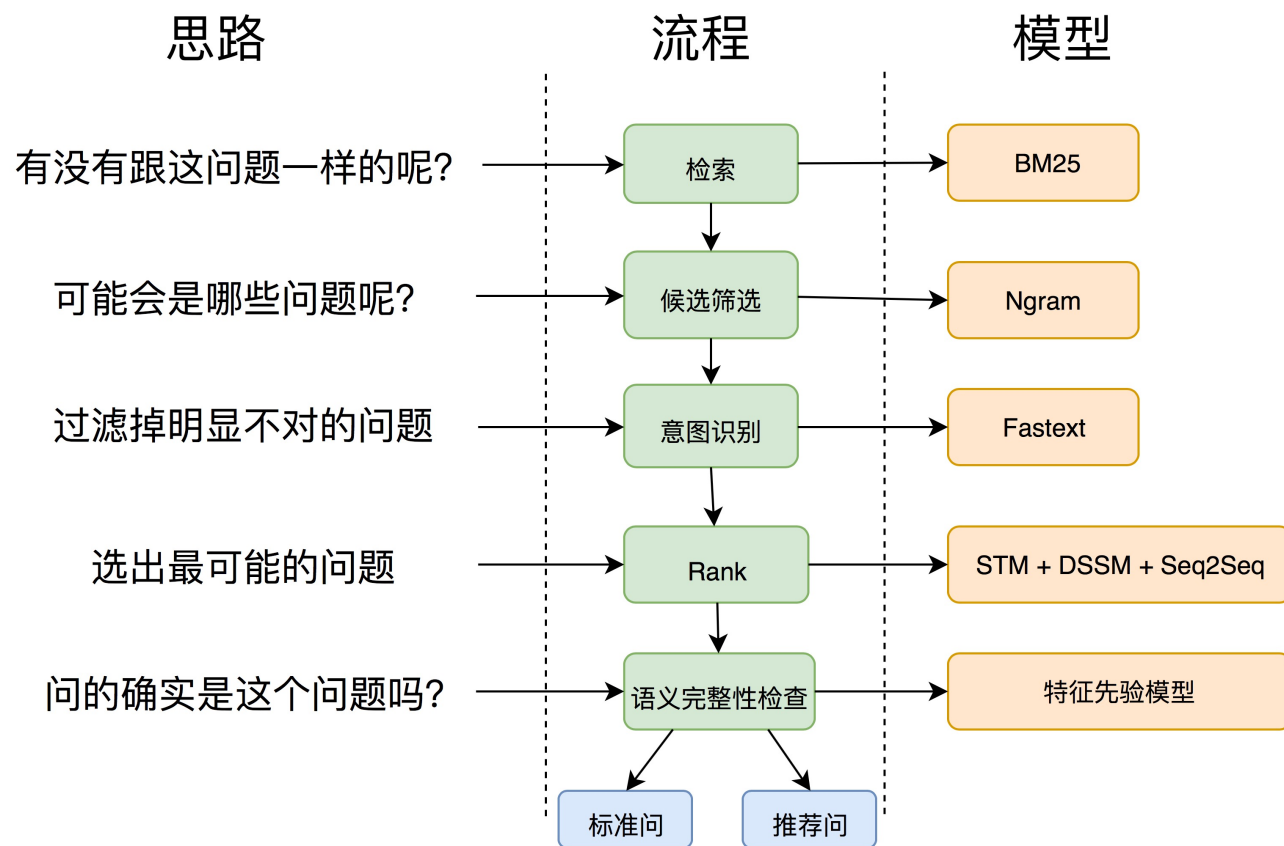
包子酸了.
吃出了两个异物,
送的腐竹是长毛的
大早上吃出来苍蝇
一杯可乐兑了半杯水
水果店给我坏的水果
这肠子怎么有点味呀
吃到一半还吃出了异物
干煸四季豆都黑糊糊了
米线里有苍蝇, 很恶心



餐品有质量问题

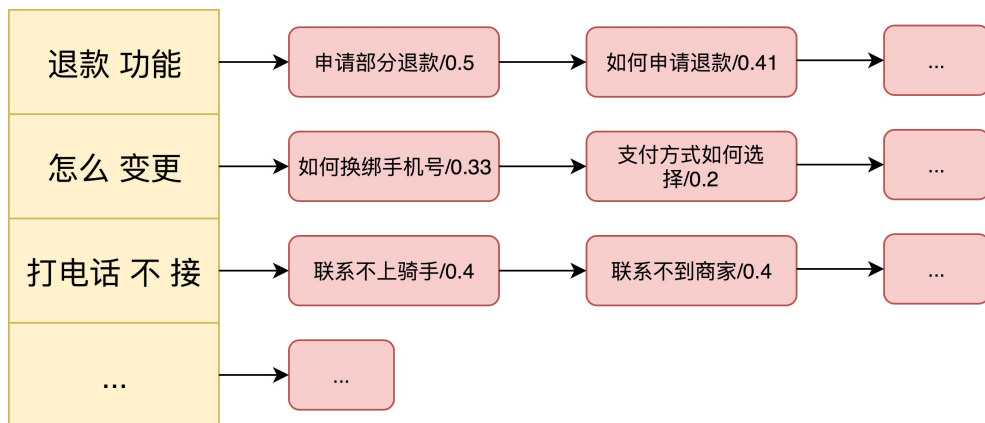
- 理解用户的问题
- 综合利用搜索技术、翻译技术、图谱技术、深度学习和统计学习技术

2.4.2 语义识别流程



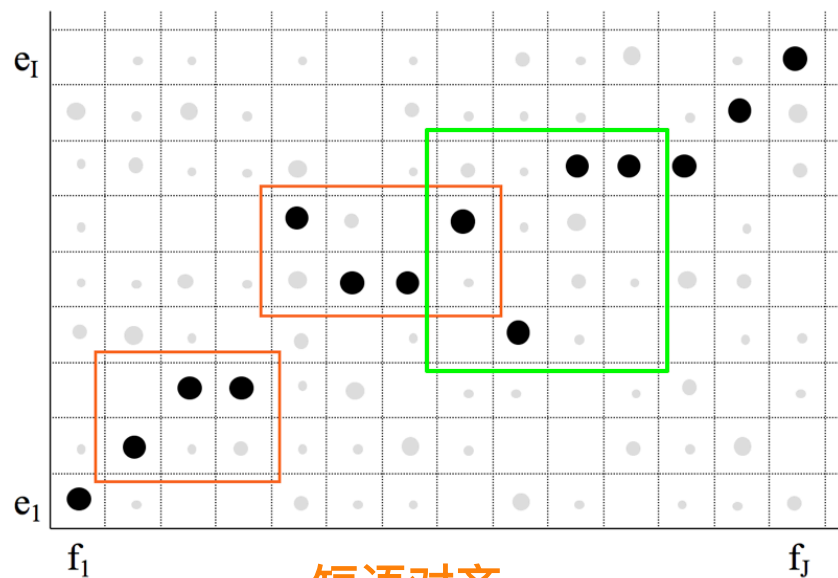
- 利用模型模拟人的处理过程
- 综合应用了搜索、机器翻译、深度学习以及统计学习相关技术
- 模型具有良好的迁移性

2.4.3 基于匹配的识别



- 拓展问与标准问的并行数据作为训练语料
- 拓展问Ngram特征串作为关键词建立倒排
- Ngram特征分布作为权重
- 用来做粗筛，Top20 准确率：99.76%

2.4.4 基于短语翻译的识别



短语对齐

$$f(y) = h(e(y)) + \sum_{k=1}^L g(p_k) + \sum_{k=1}^{L-1} \eta \times |t(p_k) + 1 - s(p_{k+1})|$$

语言模
型分数

短语对
齐分数

翻译变
形分数

$$h(e(y)) = \log \prod_{i=1}^m q(e_i | e_{i-2}, e_{i-1}) = \sum_{i=1}^m \log q(e_i | e_{i-2}, e_{i-1})$$

源语句

已超过 40 分钟 还没到

目标语句

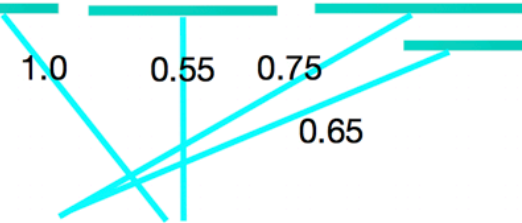
配送 超时 催单

最大值max

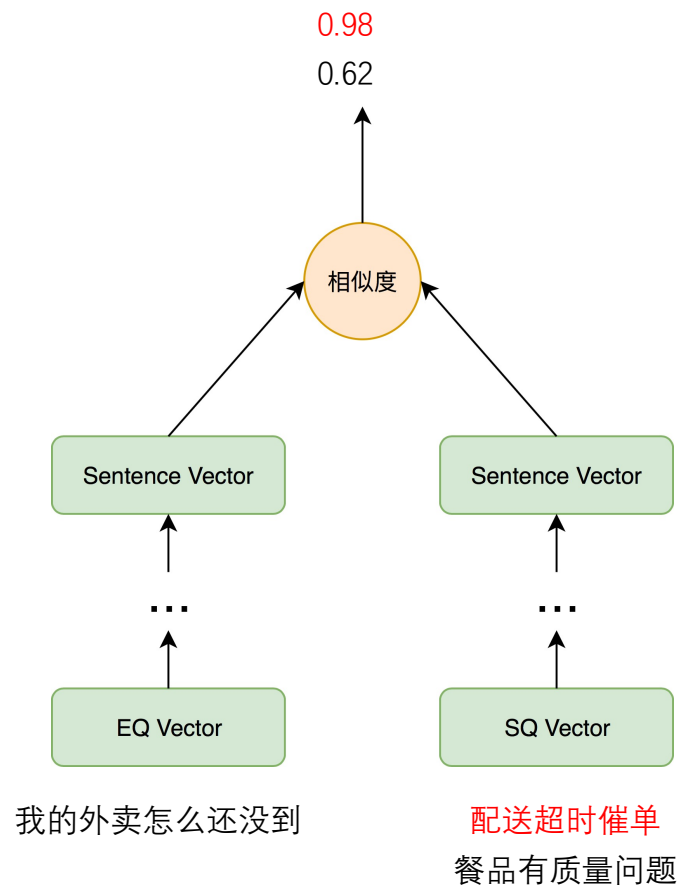
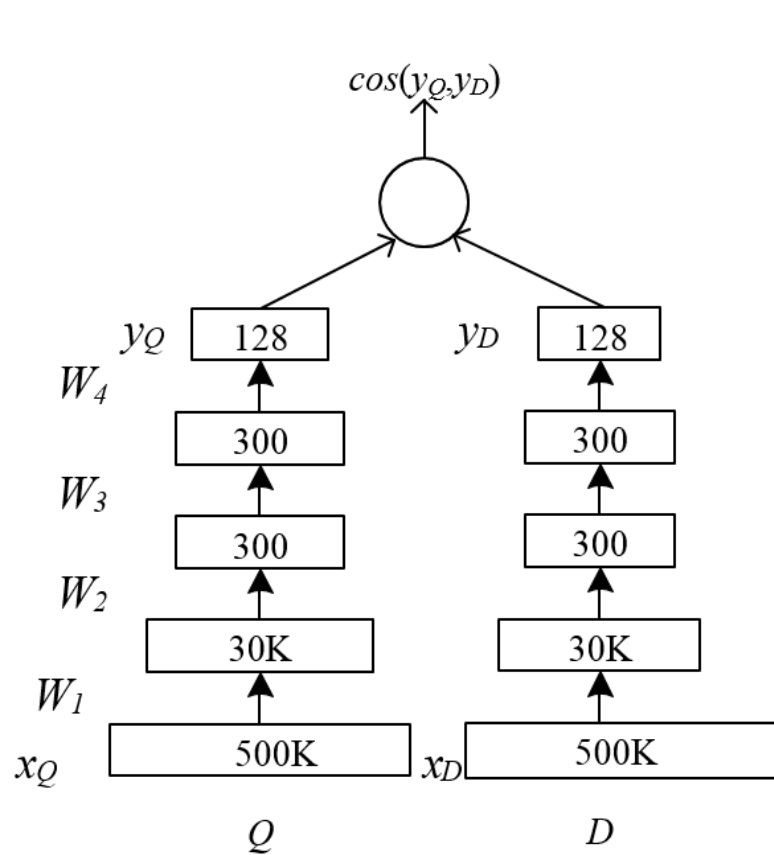
0.75 1.0

相似度

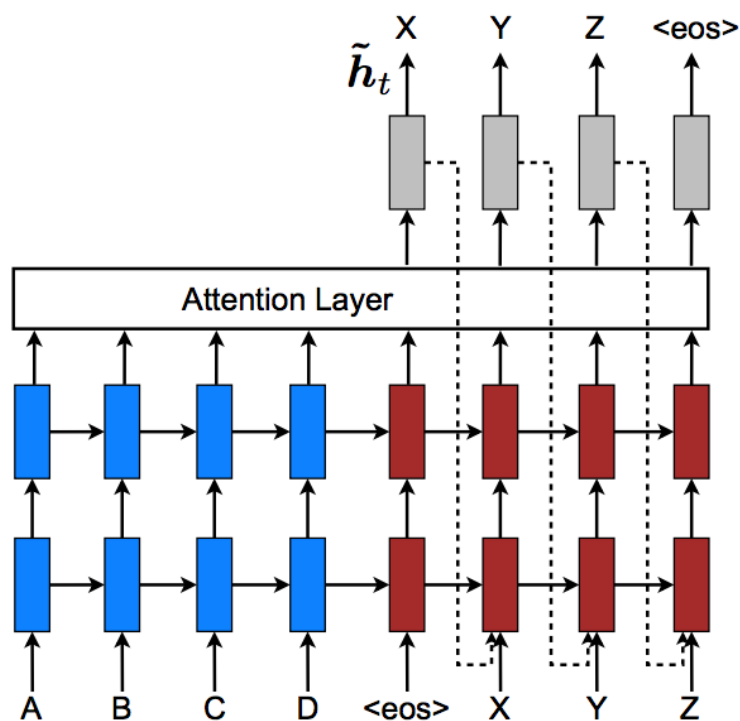
1.75



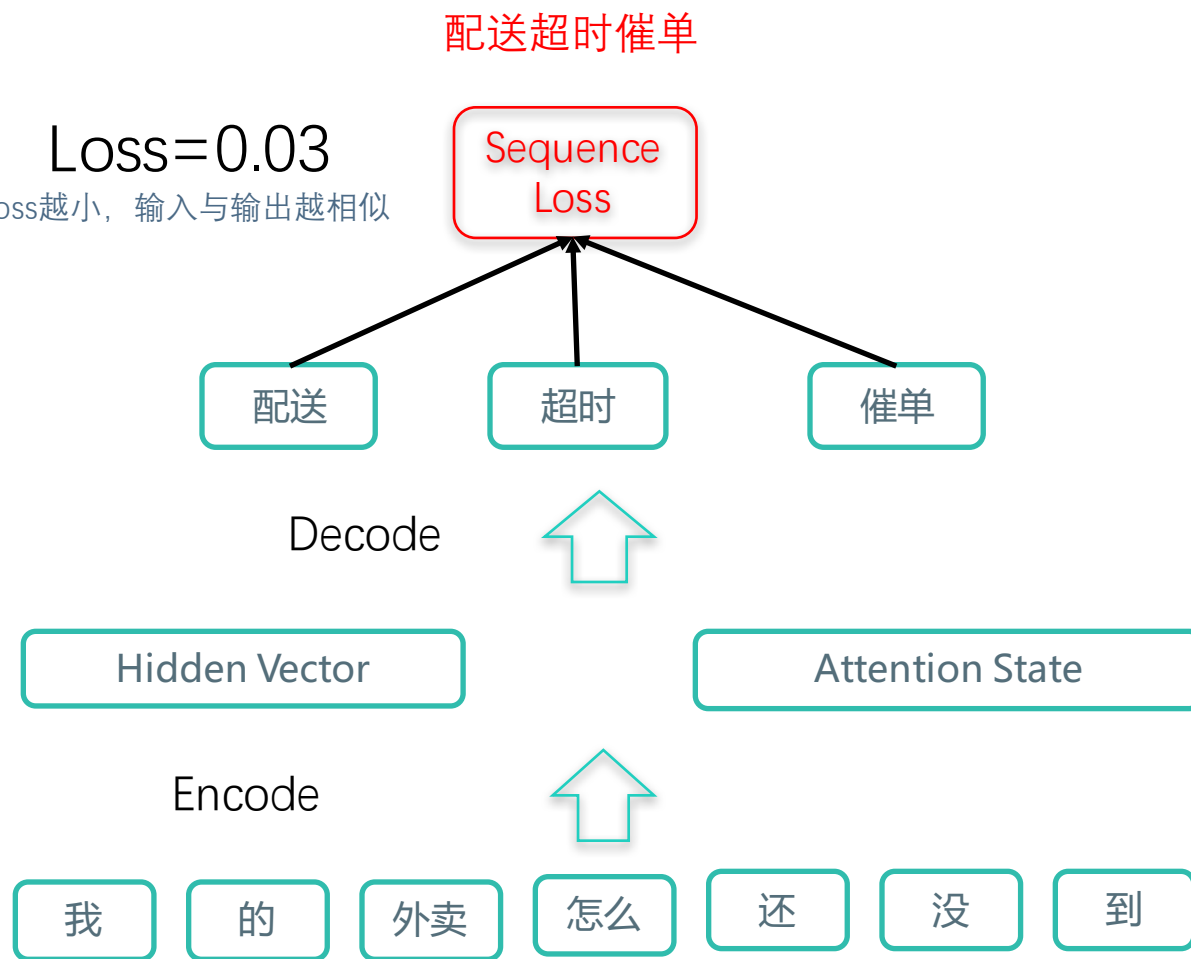
2.4.5.1 基于深度模型的识别--DSSM



2.4.5.2 基于深度模型的识别--Seq2Seq



Loss=0.03
loss越小，输入与输出越相似



2.4.6.1 基于知识图谱的识别--隐式

图谱三元组

<实体1, 实体2, 关系1>

<实体2, 实体3, 关系1>

TranE

实体向量表示

算法层

分类算法
FastText
DSSM+CNN
DNN

聚类算法
DBSCAN
KMEAN
GSDMM

特征层

GraphEmbedding

CharEmbedding

PosEmbedding

EntityEmbedding

WordEmbedding

显式特征

2.4.6.2 基于知识图谱的识别--显式

并行语料：
回锅肉 菜品
代金券 红包

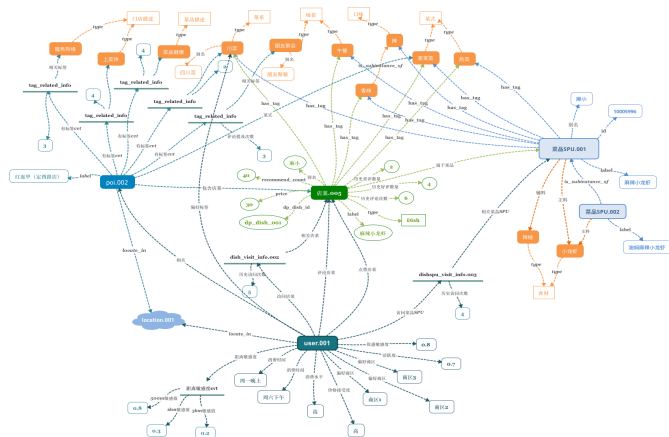
短语翻译模型

Seq2Seq

DSSM

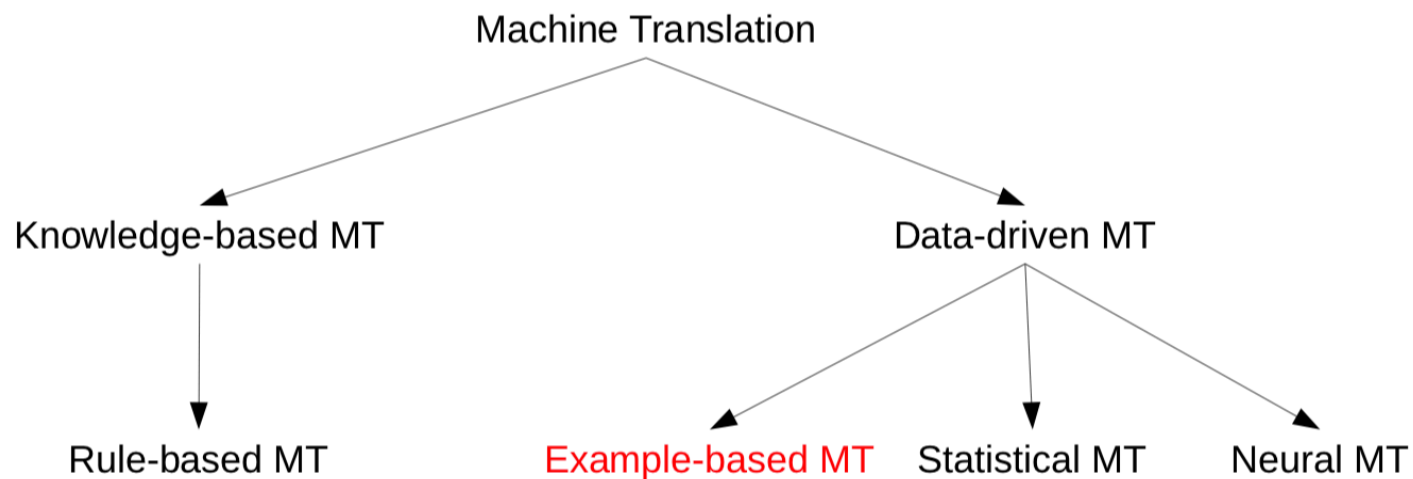
...

基于词典语义
相似度计算

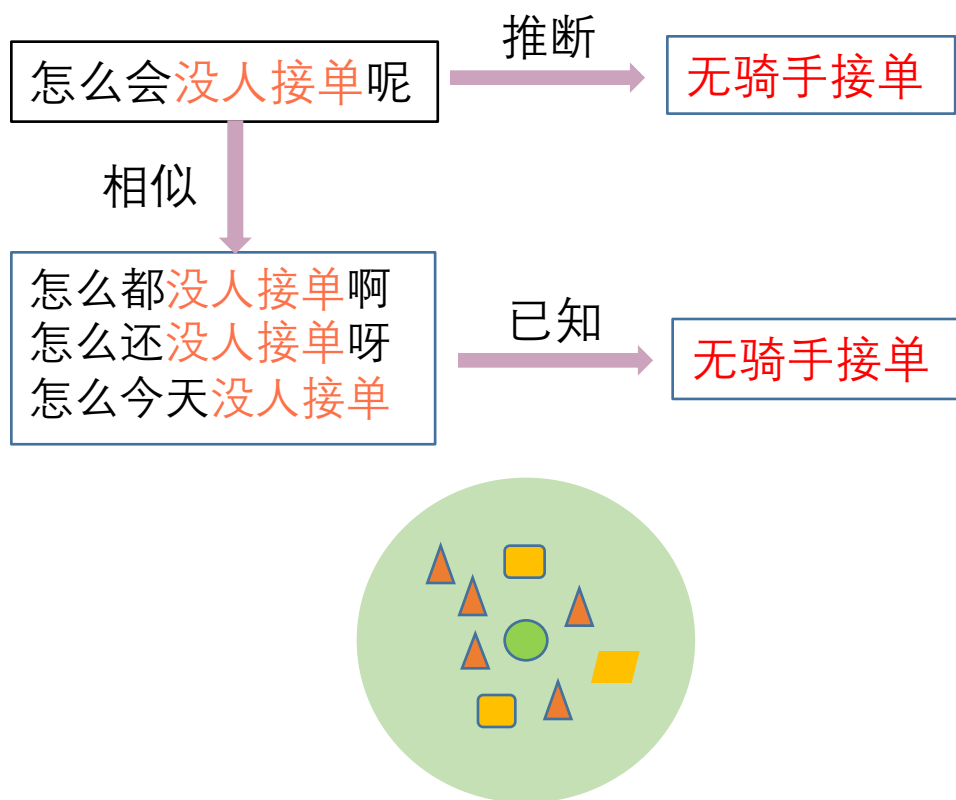


2.4.7.1 基于实例的翻译

Man does not translate a simple sentence by doing deep linguistic analysis.
--Nagao 1984

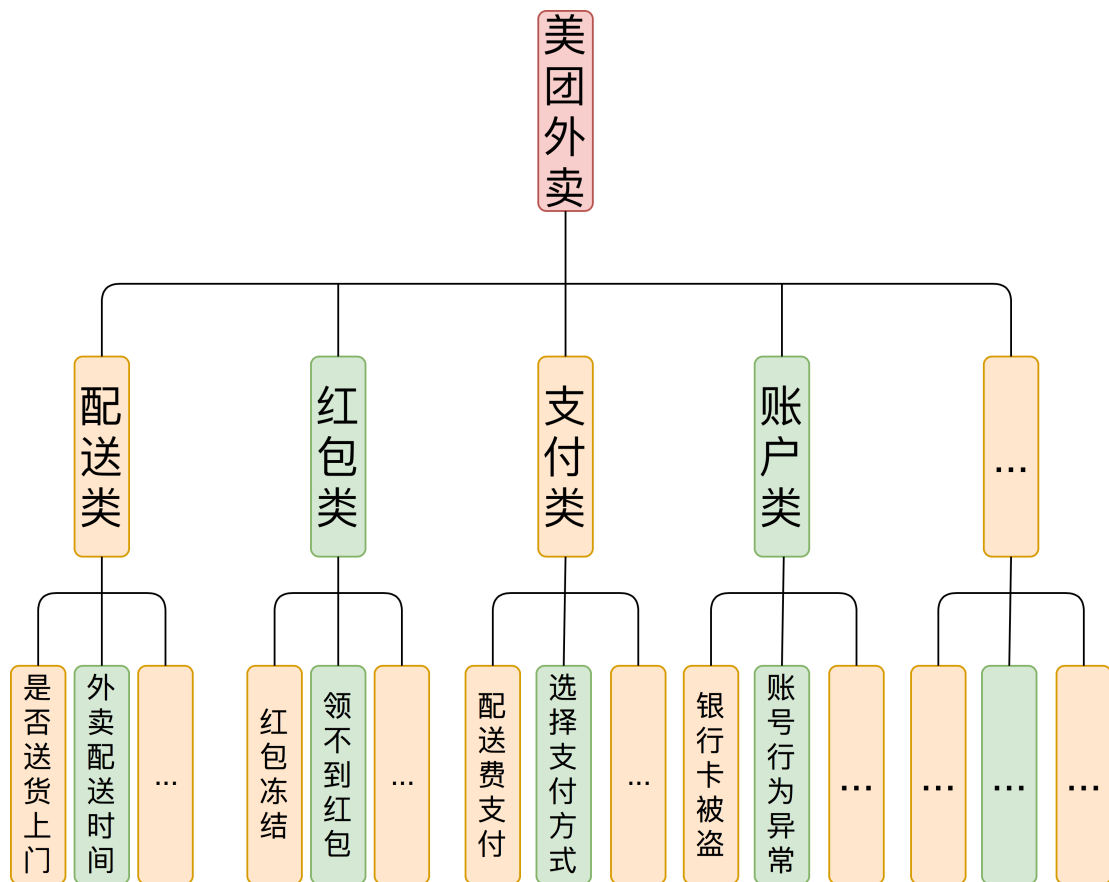


2.4.7.2 基于实例的识别



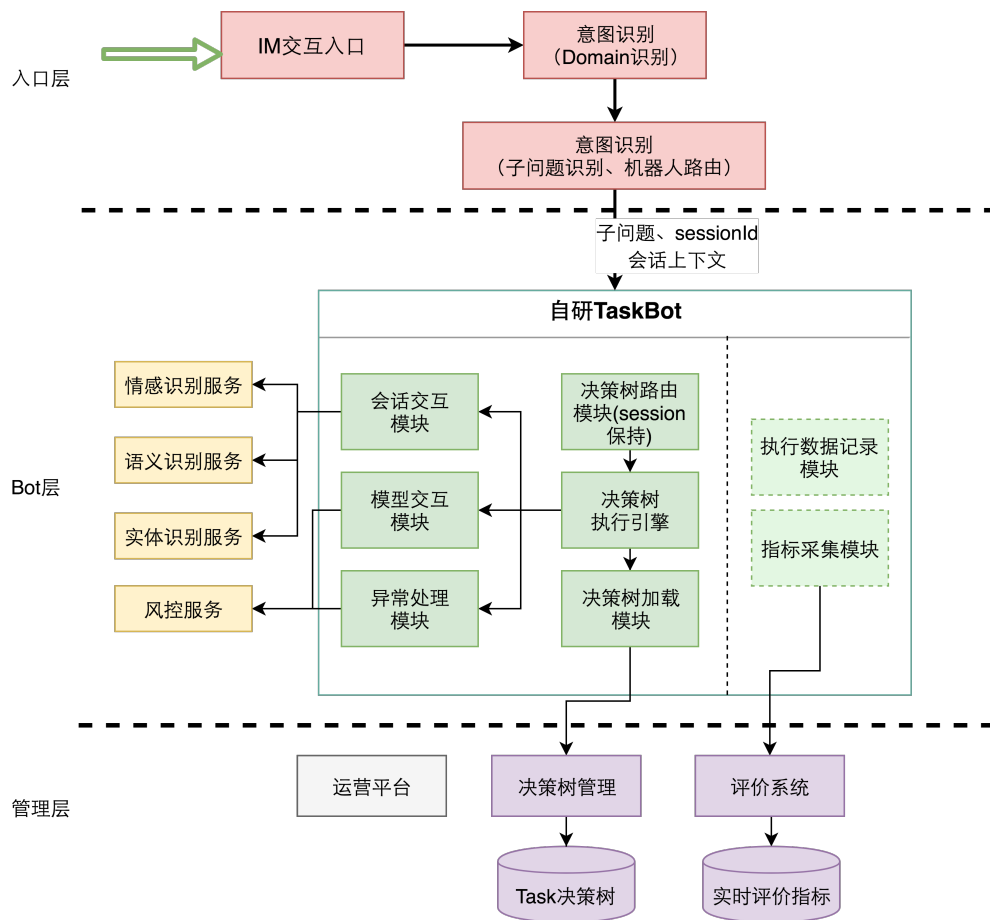
基于片断拼凑和句法分析计算相似度
特定相似度邻域内的样本进行投票决定待预测问题的答案

2.4.8 基于业务模型的意图过滤



- 根据业务逻辑建模
- 根据业务模型建立分层意图体系
- 建立意图间相容相斥关系
- 训练意图识别模型
- 根据意图识别结果和意图关系对候选进行过滤
- 保持问题的意图一致性

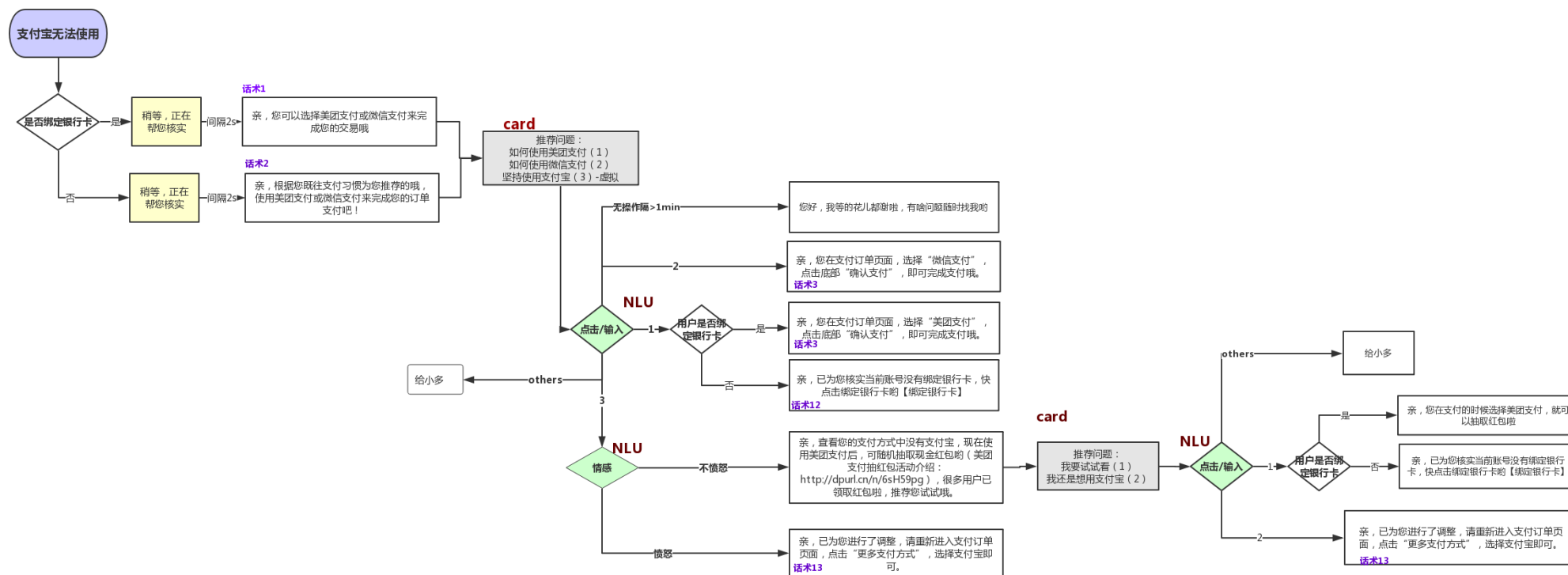
2.5 TaskBot整体框架



- 初级语义识别能力，支持反问明确用户意图
- 负面情绪判断，及时安抚或转人工处理
- 支持Task流程重入、流程回溯，SlotMap 全局共享
- Activiti二次封装：NLP和情感识别自动嵌入；error handler统一处理；存储优化、并发优化；
- HA：业务接口自动熔断；流程状态故障迁移；session sticky负载均衡 等

2.5.1 Task示例

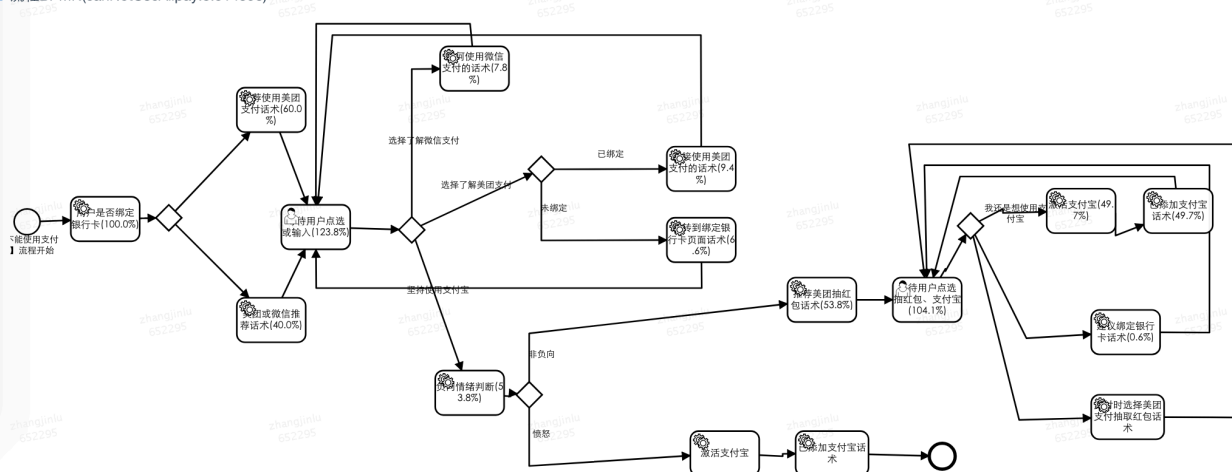
- 人工处理SOP总结task流程
- 采用BPM流程建模



Age Group	Male (%)	Female (%)
18-24	10	15
25-34	20	25
35-44	25	20
45-54	20	15
55-64	15	10
65-74	10	10
75+	10	5

c5a3c963-738d-11e8-92f2-186590df0cc5	推荐美团抽红包话术	172	0	1	否
c5a3cd91-738d-11e8-bda9-186590df0cc5	已添加支付宝话术	159	3	0	否
c5a3cb8a-738d-11e8-b51c-186590df0cc5	建议绑定银行卡话术	2	0	0	否
Task_1gcatlm	已添加支付宝话术	0	0	0	是
Task_02zij9d	支付时选择美团支付抽取红包话术	0	0	0	否
Task_0dsvf56	激活支付宝	0	0	0	否
Task_00l3sz8	激活支付宝	159	0	0	否

流程BPMN(canNotUseAlipay:3:514398)



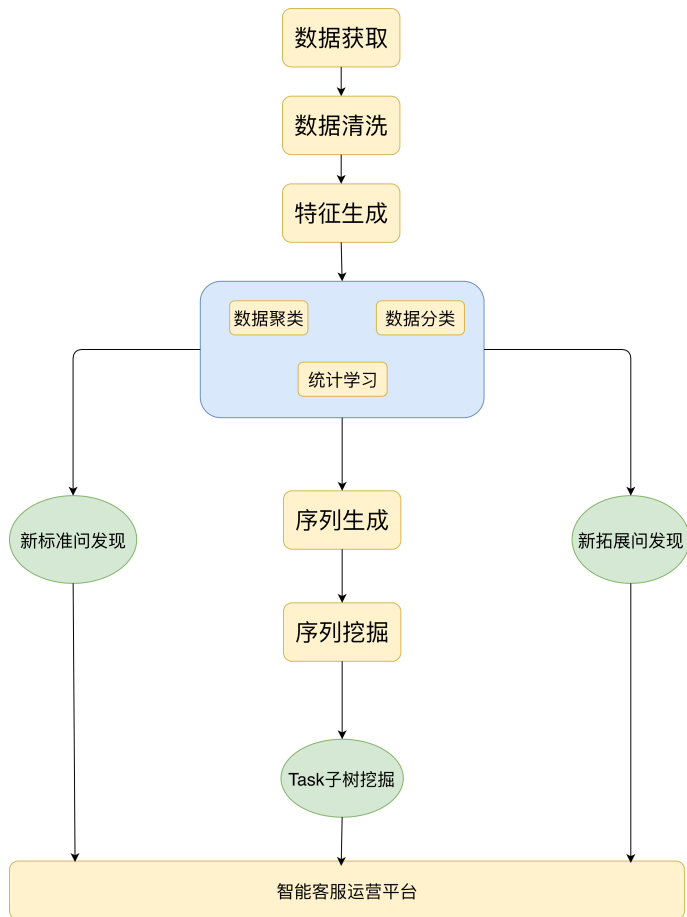
- 定义常用子流程便于共享
- 支持流程间的任意组合
- 从日志中自主学习各种子流程
- 流程与业务接口打通，实现无人自助服务
- 智能运营，不断进化

2.6 闲聊机器人ChatBot



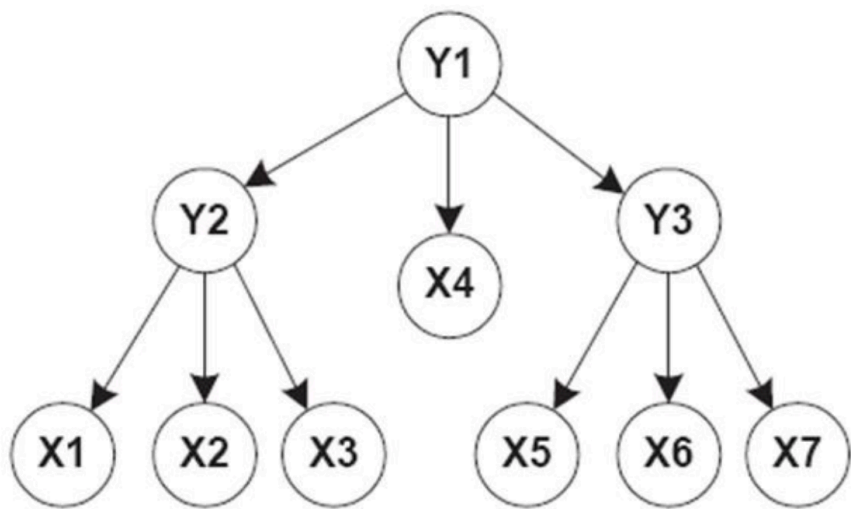
- 用户交流情感，不以解决实际问题为目的
- 主动问客户问题，收集信息以提供更好服务
- 不同客户机器人间切换的平滑剂
- 检索式：构建一个闲聊库，检索给出答案
- 生成式：从闲聊库学习生成模型

2.7 自主学习流程



- 建立统一的离线学习流程
- 从人工客服日志中学习人的处理方式
- 算法产生候选，人工轻度干预，加速模型进化
- 前期聚焦于新标准问发现、新拓展问发现、关联问题挖掘、Task子树挖掘
- 与知识管理平台打通，通过知识管理平台进行知识管理
- 关联问题实现伪多轮会话
- Task子树自动构建 + 会话管理 = 多轮会话

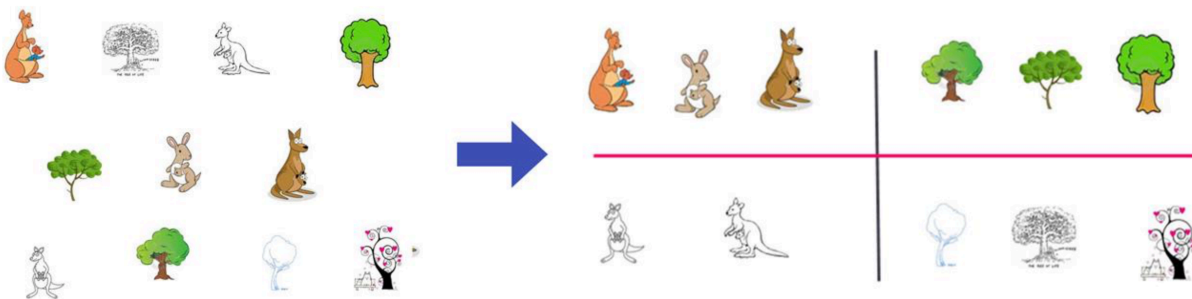
2.7.1 基于LTM的多维聚类



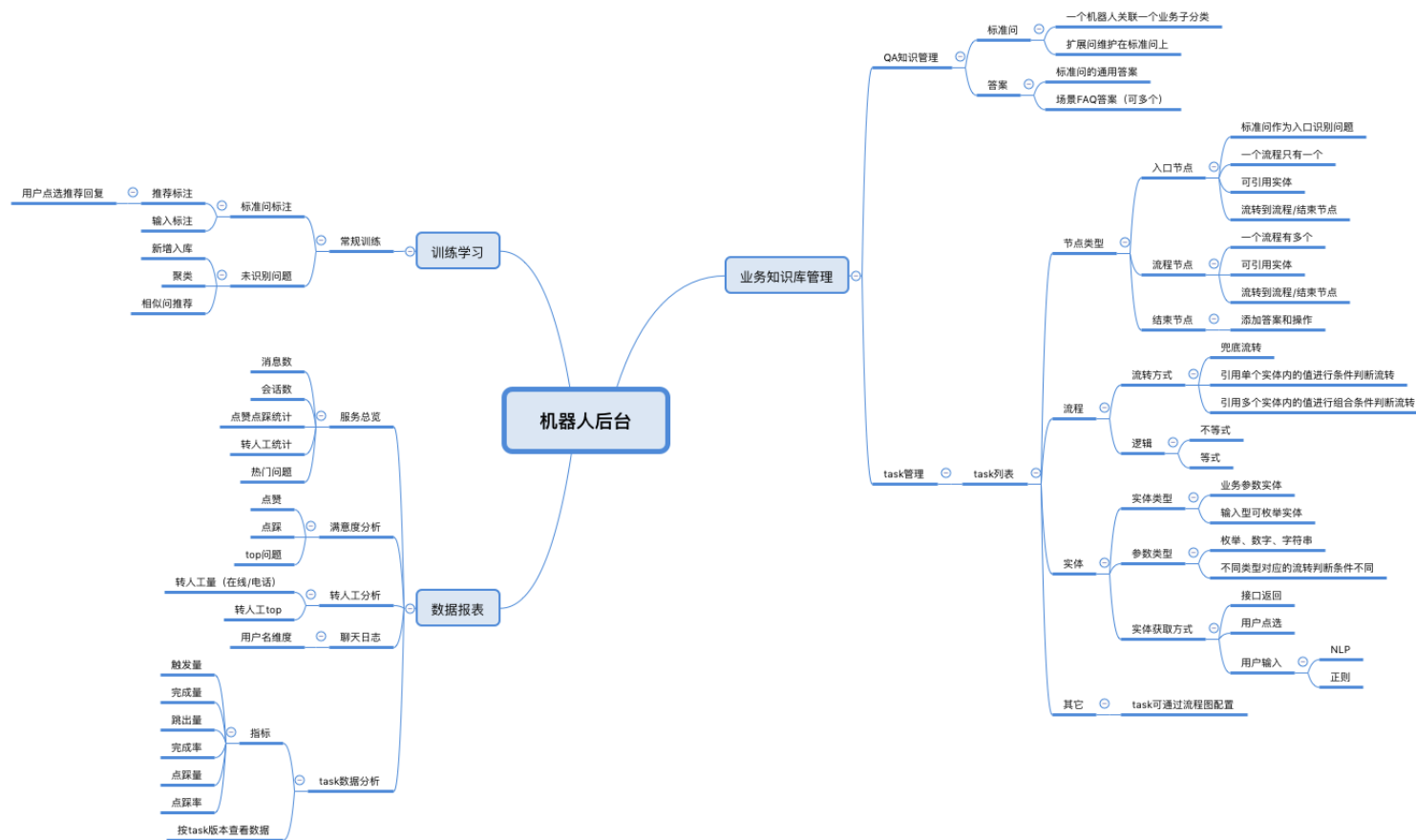
➤ LTM : latent tree model

离散变量的树形Bayesian网络

- 可以进行多维度分类，不同的离散变量代表不同的划分方式
- 无监督算法，不需要标注数据
- 可以从不同的切面对数据进行深入分析，辅助人工建模



2.7.2 人机协同学习的智能运营



AI训练师 + 智能客服 = 很多很多的人工客服

3.1 落地效果--QABot

实时响应、24小时永不离线、轻松应对高峰咨询量、比人工更有耐心



线上
效果
(外卖)

每天解决**72,000**个问题

≈ **510**个客服努力工作一天

≈ **4,100**小时

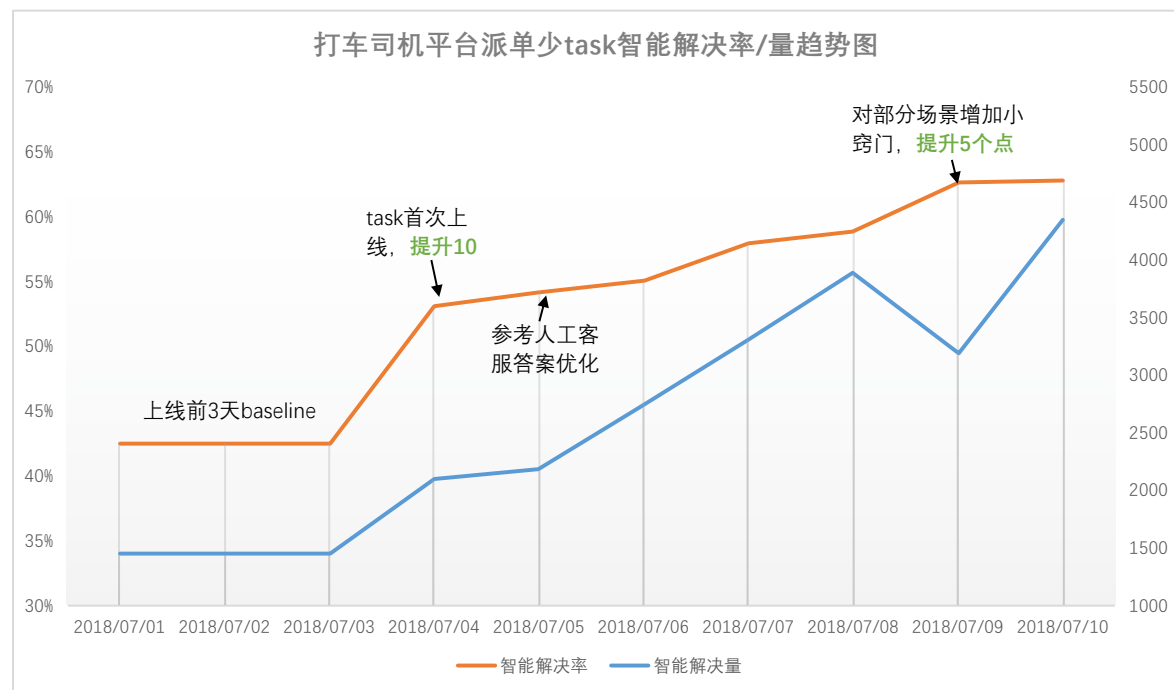
算法离线准确率 92%

在线智能解决率 83%

3.2 落地效果--TaskBot



多轮交互能力 + 情感识别能力, 为用户量身计算答案



4.总结

趋势

- 智能化是客服系统发展的趋势
- 智能客服是解决有限的客服资源与不断增长的海量用户服务请求矛盾的唯一选择

省力

- 智能客服可以大量消除人的重复劳动
- 业务专家从繁杂中解脱出来可以有更多的时间进行方案优化

协同

- 人与机器相互协助，各用所长
- 人机协同学习，教会机器，它可以帮你更多

进化

- 从日志中自主学习，智能化运营,可显著提升人客服人效

全程

- 智能客服不仅可提供售后服务，还可用于消费全流程。



Q&A

Thanks

既往不恋，纵情向前