



智能问答系统的落地和启发



徐富盛



目录 CONTENTS



项目背景和问题分析

技术架构和处理流程

实验结果

一些思考和展望



项目背景和问题分析

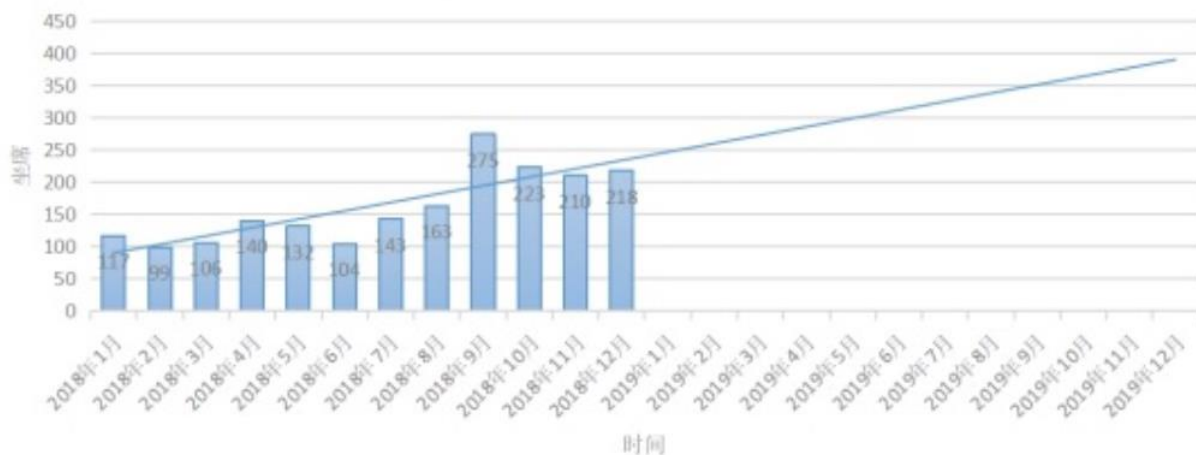
阐述现有痛点，以及相关问题的分析。



项目背景和问题分析

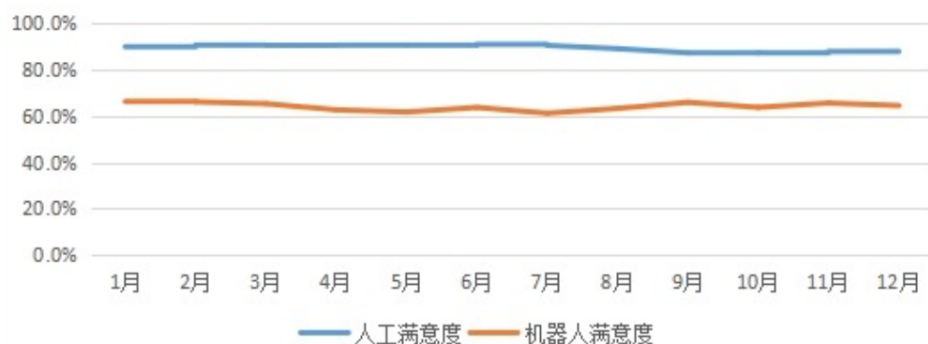
现状:

18年月在线坐席数和19年预测



- 在线客服团队人数：200余人
- 年人力成本：2000万元左右

18年每月机器人满意度



现状存在的问题:

1. 人工服务虽然客户满意度较高，但是平均每日的服务量有限
2. 如果一直保持人工服务，依照当前的业务增长速度，预计到2019年末年人力成本将达到4000万元左右





难点:

- 税务术语繁多，一般的业务员很难记得全；
- 税务相关的政策多而复杂，业务员难以准确定位；
- 税务词汇晦涩难懂，咨询者难以准确咨询；
- 晚上咨询没有业务员在线，无法及时解决咨询问题等
-



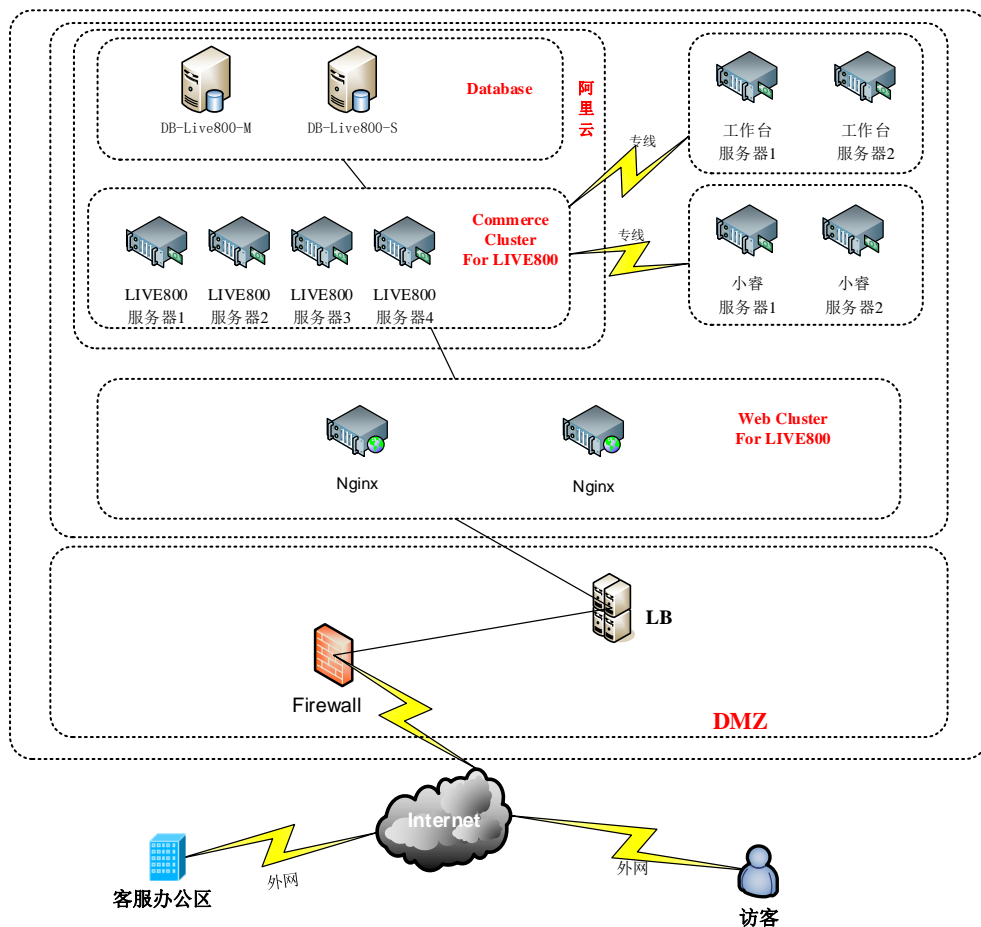
•项目背景和问题分析

- 在税务操作类问题中，据不完全统计有80%以上是重复性问题，具有一定智能的问答机器人应该能满足这部分需求。之所以现有系统并未用机器人来完成该类问题，主要原因有两个：
 1. 现使用的小i机器人问题的模糊匹配准确率只有15%左右，导致当用户使用机器人时的非常满意度只有65%左右，和人工服务差距较大。
 2. 现没有有效手段来区分重复性问题和困难问题，并且在系统层面进行人与机器人的无缝切换。
- 新的机器人问答项目将解决以上两个问题，使得问答机器人可以覆盖绝大部分的税务操作类咨询业务，从而减少在线客服的人力需求，大幅度降低运营成本和提高服务效率。





现有的技术架构解决方案

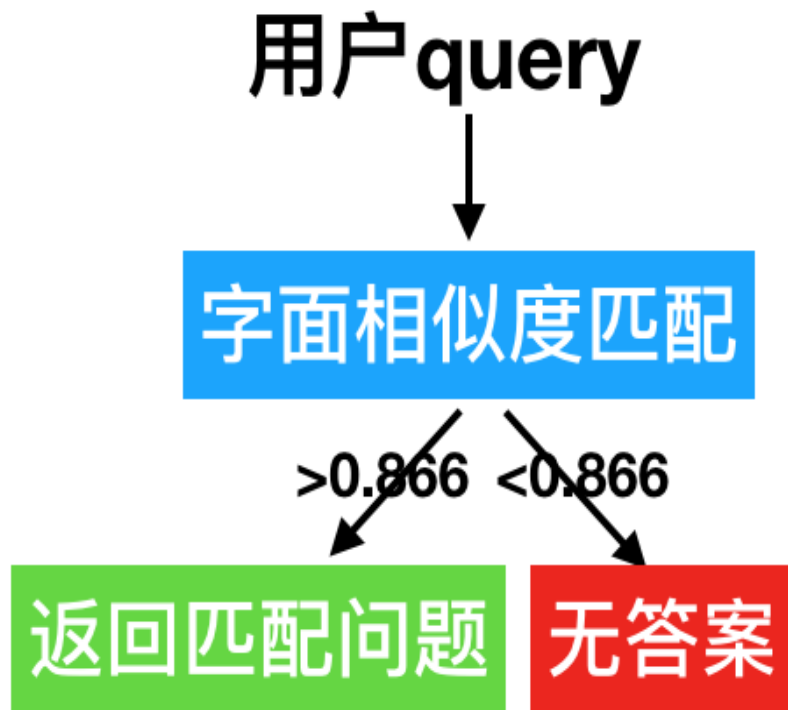


- 接入系统：自建系统，负责将各渠道方的咨询组装之后重定向到live800
- Live800系统：人工咨询，与小睿、税友深度集成
- 小机器人：机器人服务，维护问题，根据问法推荐答案





小i机器人案例：



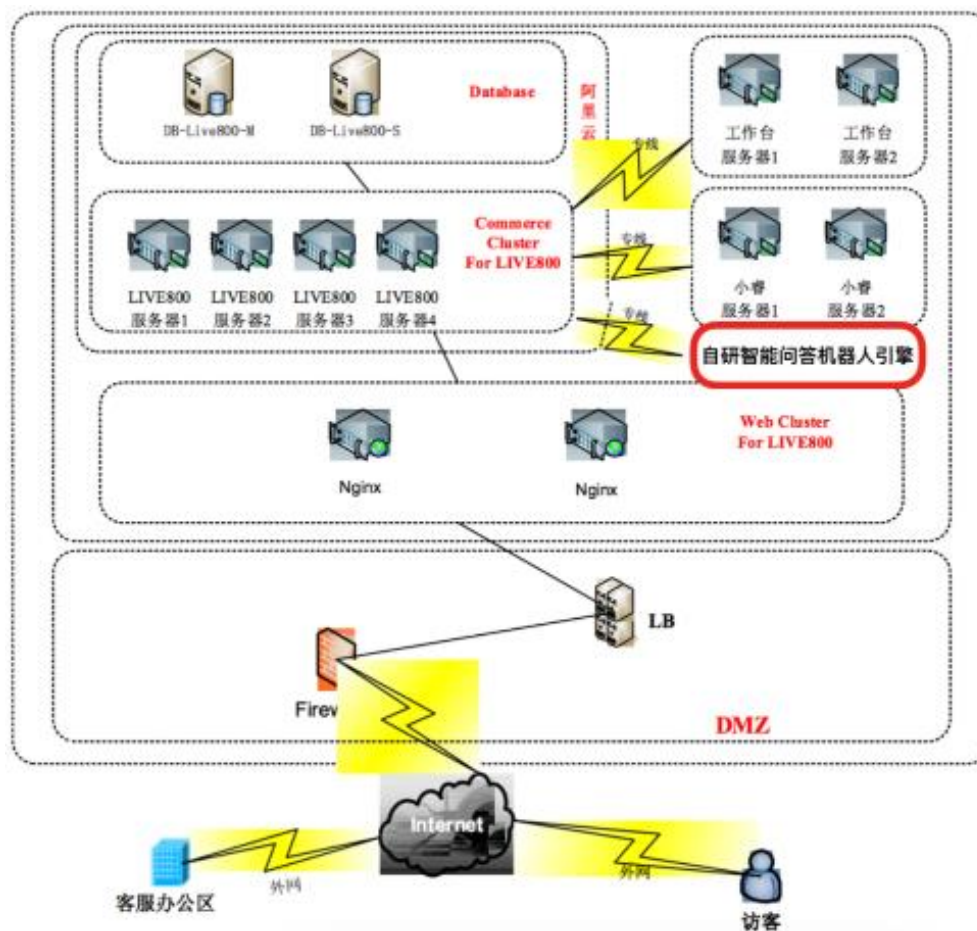
存在问题：

- 硬性的相似度阈值设定，无法识别用词相近却语义相差较大的个例，智能化水平偏低。
- 不具有基于语义的模糊匹配的功能，对于近义词、同义词等需要通过人工设置来完成。
- 基于规则算法而不是机器学习，不具有自我进化功能，较难对题库进行持续性的优化。
- 需要耗费大量人力维护扩展问。



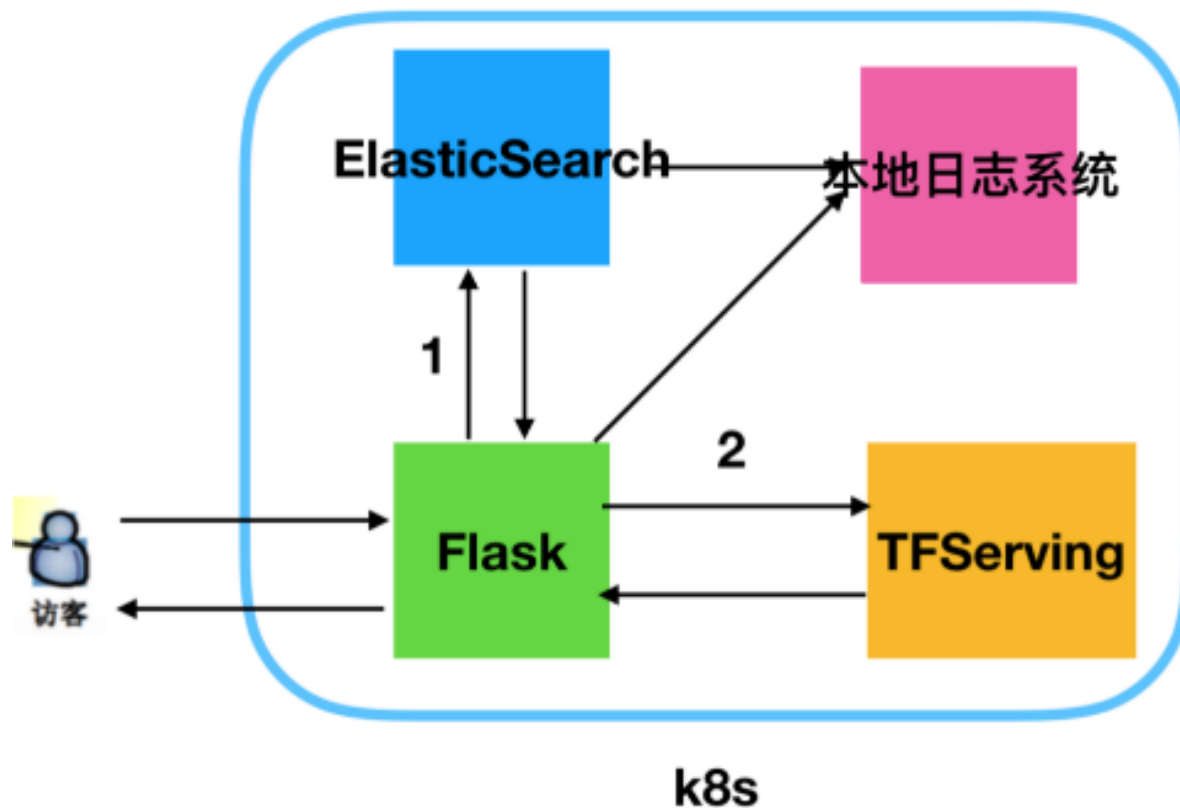


新机器人解决方案



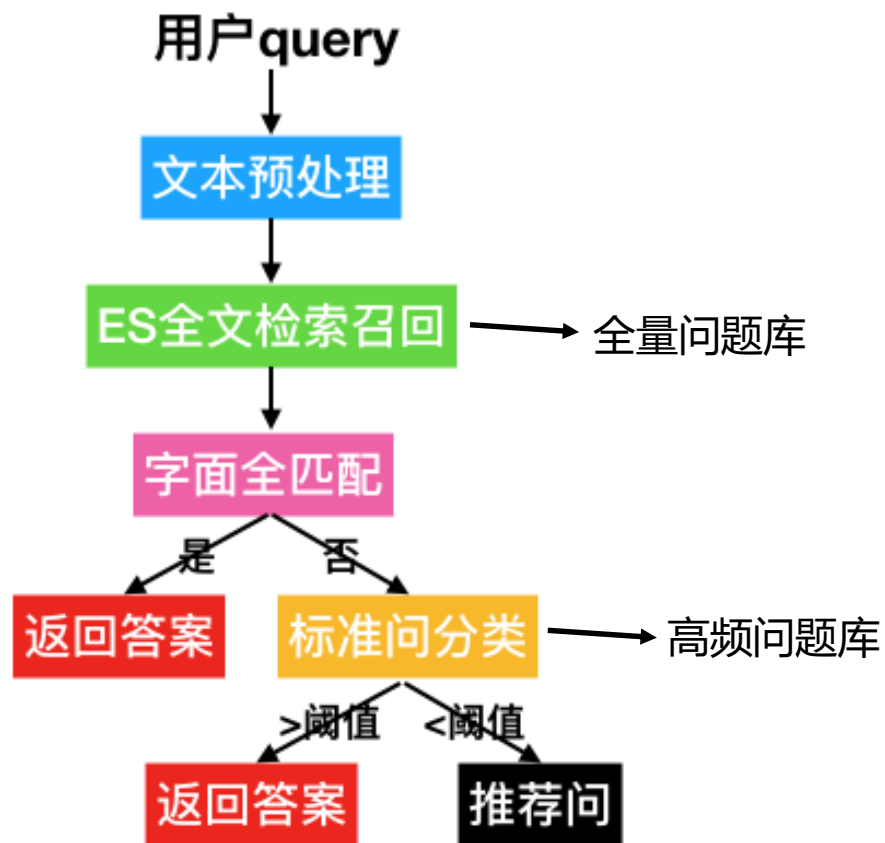


技术架构





服务流程



•技术架构和处理流程

高频问题库获取

拉取服务日志

↓
阈值筛选

↓
分组聚合

↓
人工精筛

↓
增量训练

1. 分析新机器人日志，筛选出评分（预测置信度）小于0.9的所有记录。
2. 将筛选结果按照模型计算结果进行分组，并统计每组频率。
3. 取频率最高的前10组数据，每组数据为一个聚簇（高频新问题簇）。
4. 聚簇内人工稍加筛选即可得到高频新问题。
5. 加入高频新问题聚簇数据重训练模型。



•实验结果



评测时间	测试数据 条数	平均响应时长 (ms) 小i - 新机器人	正确率 小i - 新机器人	备注
2019-03-12	200	-	62.0% - 67.7%	评测数据存在非议
2019-03-14	201	-	1.8% - 79.7%	北京自然人个税
2019-03-15	200	-	20.3% - 69.0%	北京标准版个税
2019-03-25	250	95.26 - 16.9	25.9% - 63.1%	
2019-03-29	334	83.77 - 15.0	34.5% - 78.0%	



一些思考和展望

存在的缺点：

- 容易受到对抗样本的攻击，使得精度下降；
- 拓展问不够多
- 无法实现多轮交互对话
- 无多模态的数据来源
- 。 。 。

思考：

- 通过扰动来获取更多的样本，从而降低对抗样本的干扰；
- 建立领域的知识图谱，扩展相对应的关联问题；
- 需要从用户问题中挖掘出全部多轮任务型问答；
- 。 。 。





THANK YOU



xf665486