

UNIT 项目汇报

6/13/2018

1. 问题

- 1) 能否将多个技能进行组合，构成一个更复杂的 BOT？
- 2) 如何避免直接回复『我不知道应该怎么答复您』而是向用户提供可能的意图？

2. 更新：现在可以使用日志分析功能了。日志记录了用户和 BOT 的交互，开发者可以了解哪些用户 query 没能被识别出意图，之后对 BOT 进行改进。

UNIT 平台使用

UNIT 平台提供基于深度学习的**自然语言文本解析服务**和**对话系统**。开发者可以训练一个具有**特定技能**的机器人（BOT）。该 BOT 接收自然语言文本（可以是口语化的表达方式），通过和用户的交互式问答，解析出用户的意图，给出相应答复。答复可以是文本，也可以是开发者指定的执行函数。

一个简单案例

以微信小程序举例：利用百度语音、UNIT，实现一个可以根据用户要求，播放歌曲的微信小程序『点歌服务』。

Step1

用户进入『点歌服务』，录制了一段语音来描述 ta 想点的歌。

Step2

『点歌服务』调用百度语音提供的接口，将这段语音解析成文本。

Step3

『点歌服务』将文本上传到 UNIT。

Step4

我们在 UNIT 平台上训练好的 BOT 会对文本进行解析，识别文本中的用户意图。假设所有 BOT 回应的触发条件已经满足，那么 BOT 给出回应，并返回给『点歌服务』。

Step5

『点歌服务』得到回应。根据其中的执行函数和词槽调用相应资源，为用户播放 ta 想要听的歌曲。

BOT 训练流程

1. 新建一个 BOT

注意：

- 1) 一个 BOT 只能对应一个技能

2. 设定意图

意图分为『对话意图』和『问答意图』。对话意图用于较复杂的场景，可以解析口语化的自然文本。问答意图用于简单的场景，如 FAQ。

两种意图中，主要开发更为强大的对话意图。为了方便叙述，后面以『意图』简代『对话意图』

首先『意图』与『词槽』相关联。

『词槽』是一种词典，里面的词属于同一种类别，比如天气可以是一个词槽，其中包含晴天，阴天，暴雨等描述不同天气的词。『词槽』往往是查询『意图』的条件。

eg: 『意图』：查询是否下雨。而为了查询下雨情况，需要知道时间和地点这两个条件来组成一个完整的问题，如：

周六北京下雨吗？

由于这两个条件是查询天气必要的，而且每个都是一类词，那么就要把时间和地点设定为两个词槽。

之后为『意图』设定回应方式，可以是一段文本或执行函数（『答复』），也可以是『引导至对话意图』或『引导至问答意图』。

回应要有触发条件。在已经识别了一个意图的情况下，如果回应的触发条件没有达到，缺少『词槽』，那么 BOT 会要求用户进行澄清。

3. 设定对话样本集合

『对话样本』是用户可能给出的实际对话文本，用来训练 BOT。

需要对样本进行标注，这包括：选择意图和标注词槽。只有确认被标注的文本才可以用于训练。

4. 设定对话模板集合和特征词

『对话模板』是可以被 BOT 识别的对话规则。

每个模板对应一个意图。

模板由词槽、特征词、一般文本组成。

- 可以设定词槽和特征词的出现顺序和是否必须匹配。
- 可以设定模板被识别的阈值（文本中可识别部分长度与文本总长的比值）

『特征词』是某些具有相同特征的词的集合。在用户给出的 query 文本中，会有某些词经常出现而且代表的是相同的含义，那么就把这些词收集并创建一个特征词列表，在模板中使用。更多的『特征词』有利于提升 BOT 能力。

5. 训练模型并体验和 BOT 的对话

略