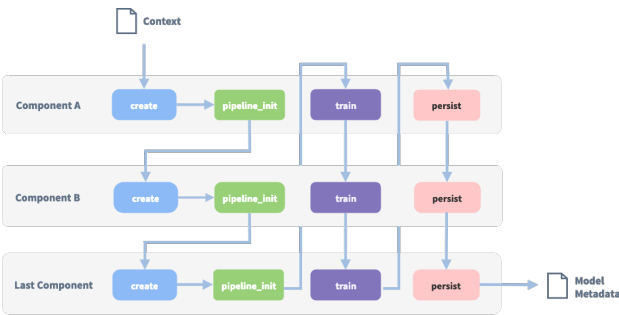
关于rasa 中的源代码的一些心得

NLU

这部分类似管道的概念



将context一步一步经过多个component部分，然后再将结果拿去train。

整个component的配置在于config.yml中，这其中又包括了分词，特征提取信息，分类以及实体识别。这部分类似做一个特征工程，提取出输入的Slot/Entity 跟intent。

分词部分包括whitespaceTokenizer、JiebaTokenizer、mitieTokenizer、spacyTokenizer等

特征提取信息包括mitieFeatures、spacyFeatures、languageModelFeatures、RegexFeatures、countvectorsFeatures等

分类包括mitieIntentclassifier、scklearnclassifier、keywordIntentclassifier、

DIETclassifier等

实体识别包括spacyEntityExtractor、CRFEntityExtractor、DIETclassifier等

当我们配置完这些时，都应该测试nlu部分，便于我们了解其中采取的策略是否合理有效。

Policies

这部分同样也是需要配置在config.yml中。

TEDPolicy

跟机器学习相关的

RulePolicy

规则相关的

MemoizationPolicy

内存记忆相关的这部分的泛化能力较弱。

Action

这部分是用来定义用户的行为的基础，例如我们用来联系第三方、查询天气、定时等等产生用户行为动作的事件都需要在Action中重写name、run方法。它会返回一个Event事件信息。

Action中包括了储存记忆的tracker，可以针对sloty采取动作，可以对action回应的dispatcher。

Events

Rasa对话都是呈现一系列事件的。Rasa SDK提供了每个Event的类。

知识库Action

当我们需要对答的是一种类似知识信息，一方面我们可以在action中重写，也可以将相应的知识书写成json形式存储进去。

个人总结

整节课来说理论过多，我们需要实践来掌握这些信息，同时也希望老师们能提供一些丰富例子供我们理解理论。