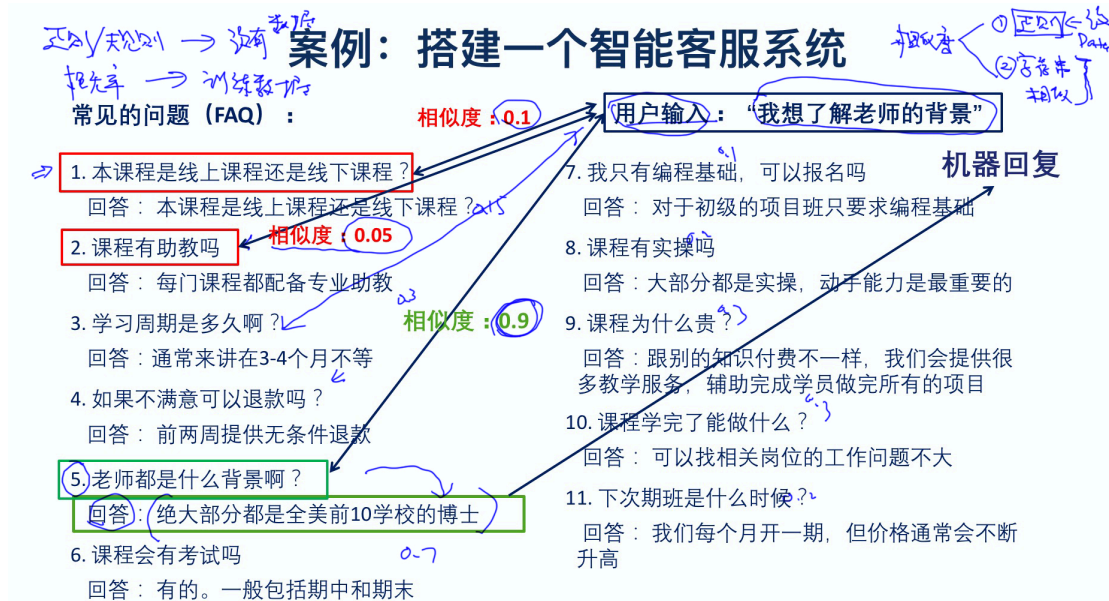


1. 搭建一个智能客服系统



保证有一定的语料库, knowledge base,

2. 基于搜索的问答系统

考虑英文的方法 (因为有些方法英文和中文不兼容)



对于问题, 分词, 预处理:

1. spell correction
2. stemming
3. stop-words
4. words filter
5. 同义词

将文本转换成向量:

1. boolean vector
2. count vector
3. tf
4. word2vec
5. seq2seq

给定两个向量，计算相似度，将相似度排序。

3.

问答系统核心：文本表示、相似度

知识图谱：实体抽取、关系抽取

输入了一句话、一个问题，需要到知识图谱里寻找答案。

贪心学院什么时候成立？

实体：贪心学院。关系：成立。

