

第六天

2019年6月14日 10:09

聊天机器人的类型

1. QA bot : 问答机器人
 - a. 信息检索, 搜索
 - b. 知识图谱
2. TASK BOT: 任务机器人, 帮人做事情
 - a. 槽位填充
3. Chat BOT: 闲聊机器人
 - a. 信息检索
 - b. seq2seq模型

项目的实现逻辑:

1. 问题进行基础的处理
2. 判断用户意图
3. 如果用户进行闲聊, 调用闲聊模型返回结果
4. 如果用户要进行问答, 调用问答模型返回结果

词典的来源:

1. 输入法的词典
2. 手动收集

收集停用词

jieba

1. jieba.lcut
2. jieba.load_userdict(path)
3. jieba.posseg.cut() #返回词性

fasttext的使用方法

1. fasttext数据格式
 - a. 分词后的句子\t_lable__label
2. import fastText
 - a. model = fastText.train_supervised(path,dim,...)
 - b. model.save_model(path)

- c. `model = fastText.load_model(path)`
- d. `([predict],array([prob])) = mode.predict(sentence)`