**MedicalChatBox 工作记录**

**20210114** **工作复现**

《Effective Medical Test Suggestions Using Deep Reinforcement Learning》

代码和数据文件暂不可知，代码文件估计要发邮件问

数据文件来自SymCat，去除了所有不在CDC database里面的症状，并将UMLS 症状库中的二级症状（归属于其他症状之下的症状）去掉

Flat-DQN 0.38-0.40

HRL 0.47-0.49

SVM-ex 0.2806

SVM-all 0.7123

**20210325工作复现**

截至目前为止，此工作的普通版本和提高运行效率的batchsize版本均调试完成，但在运行时会出现一些问题

为了更快的测试模型收敛效率，我们先对reward进行测试，按照文章中的参数在小数据集（50种疾病）上运行时，因为文章参数中提出检查建议的成本较低（约为正确预测奖励的1/100），且并未限制提出检查的数量，Symptom checker会倾向于直接进入建议医疗检查阶段，提出大量医疗检查建议，通过其差异进行判断，随后进行疾病判断，虽然正确率能达到较高水平，但在现实中显然不可行。

之后我提高了进行检查的代价并加入了要求模型重建所有异常结果的rebuild loss后，模型的对话轮次上升，对医学检查的数量要求下降，但正确率也随之下降，从90%左右下降到60%左右

之后换用200个疾病产生的100000个样本进行训练，基本情况与小数据集的结果类似；模型或者选择直接进入Test checker和Disease Classifier进行判断，要么收敛到不断对同一个症状进行query直至超出最大轮次。

原因可能是：200个疾病所表现出的症状和所需医学检查过多且彼此交叉项较少，在取样时难以获得正确的症状和疾病，且对正确预测疾病数量的奖励只有在最后一步的疾病判断中才予以体现，造成前面Symptom checker训练效果不好

综上所述，这个模型有以下可以改进的方向：

1. 从数据集上来看，应选取症状和检查相对重合度高的疾病压缩搜索空间
2. 在训练时，应给予正确预测的症状即时奖励而不是在谈话的最后算一个总数奖励
3. 为避免模型直接趋向于使用Test进行判断，应该限制提出医学检查的数量，并要求模型对所有症状和检查结果进行重建

因为包含检查结果的真实问诊数据集还没出来，所以接下来的工作有两个方向

我下一步是继续用这篇文章的policy-based方法去尝试，还是回到DQN用改进DQN样本利用率的方法去尝试？

**20210415 Baseline复现**

从周二（20210413）拿到数据之后，我在Qianlong学长的MedicalChatBox模型上进行了实验，值得注意的是，Qianlong学长的测试集采用的方法是从测试集中随机取样进行测试，而我使用的方法是遍历测试集，计算正确率，会比学长的指标略低

**实验记录**

**实验一：**

首先使用step cost = -1, correct symptom = 1，（每一个轮次的代价为-1，问询到正确症状的奖励为1），其他参数和论文中一致，网络迅速收敛到直接对疾病进行判断，相当于直接通过MLP对疾病进行判断，SR（success rate）= 0.6, AT (average turn) = 3

注：AT=3 意味着 自述——疾病判断——用户关闭对话

分析：对对话轮次进行惩罚时，在稀疏的action space里agent大多数选择只能得到负收益，所以会选择减少对话轮次，直接进行判断

**实验二：**

去掉step cost，correct symptom = 2， 在1500个epoch之后仍然收敛到直接进行疾病判断，SR = 0.586, AT = 3

分析：同实验一，仍然需要继续增加查找到正确症状/检查的reward

**实验三：**

在实验二的基础上，将correct symptom置为4 （正确判断疾病的奖励为44，错误判断为-22），这时模型的交互轮次升高，但正确率有所降低，SR = 0.34, AT=11.626。

分析：在测试集中，我们发现agent有时会选择不断query symptom而不进行disease classification，另一方面，在症状中存在重复表述的情况，如有“高热”，“地热”又有“发热”，有“感冒”也有“咳嗽”、“喷嚏”，但在病人表现出的症状中往往只有其中之一，造成request不准确

**下一步计划：**

在经过上述实验之后，我跟陈伟学长进行了讨论，认为有两件事情是现在需要完成的

1. 对症状和检查集合进行重新归一化
2. 强化学习的目标或许可以设定为“从部分症状recover所有的症状”，而不是将症状问询和疾病判断强耦合到一起完成，如果强化学习能cover所有症状的话，理论上我们便接近了预测能力的上限

所以下一步我们打算：

0、对症状和病人信息进行重新归一化

1、通过简单分类器（SVM、MLP）通过自述和完整症状计算预测能力的上下限

2、单独训练一个训练良好的agent，这个agent的目的仅仅是为了recover症状，这里可以计算一个f1

3、在test set上用1里面训练出来的分类器，去预测疾病，这里可以计算一个疾病准确率

上面2和3的指标，是二阶段的指标

4、把1和2联合起来，这样就是联合训练，会有一个联合的指标