方法

我们用了最简单的方法实现了baseline版本，得到的分数为0.68。方法详见如下所述

1. 把角色名称与对话合并成一条文本，对角色的情绪识别变成对文本分类问题，但由于有多个情绪需要识别，因此是多标签分类问题。每个标签的类别是[0,1,2,3]四种取值，但是从数据分布来看，类别2和3的占比非常少，为了简化，我们采用多标签二分类的方法。
2. 多标签分类和多分类的区别在于：
   1. 多标签分类是对一个样本的多个方面，或者多个目标进行分类，通常是二分类。这些方面不是互斥的，比如对于一张人物图片样本，在性别方面区分是男是女，在年龄方面区分是大人还是小孩，在种族方面区分是东亚人还是西方人。图片的这几个方面可以同时出现，既有性别，又有年龄，还有种族
   2. 多分类问题是对一个样本判断属于哪个类别，类别的取值超过2个。这些类别是互斥的，属于类别1就不可能属于类别2或类别3。比如对人脸图片样本，区别是谁的人脸。
3. 我们的任务是个多标签分类，因为对于一句文本，要从多个不互斥的方面（爱，乐，惊，怒，恐，哀）进行分类，为了简化，我们仅仅分类为0类或1类，而类别2和类别3当成类1来看待。
4. 把文本转成向量，我们采用目前比较流行且效果比较好的bert-base向量
   1. 我们取一批文本样本，批大小为8，算出这批数据的最大文本长度max\_length\_batch,把文本的每个字符先转成字符id，把一批长短不一的文本padding到固定长度max\_length\_batch的序列。Padding值为0
   2. 将样本的字符id序列送入bert模型，从bert模型最后一层的第0个单元取出向量作为整个句子的向量。
   3. 把得到的句向量送入全连接的网络进行分类，由于是6个标签，输出是6个神经单元，得到6个logit。
   4. 与多分类不同，多分类任务是把6个logit做softmax，而多标签分类任务是把6个logit一一做sigmoid，每个用sigmoid激活后的值代表这个标签为类别1的概率
   5. 把概率大于0.5的标签，转为1，小于等于0.5的标签转为0

