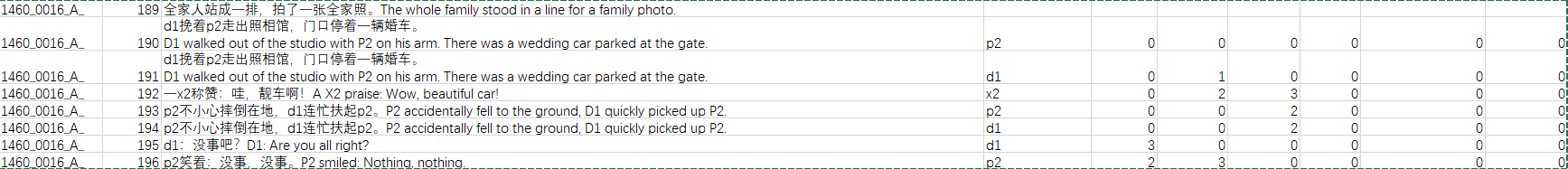
摘要：

This task is to analyze and identify the emotions of each character involved in every dialogue and action description in the script scenes from multiple dimensions. Comparing with traditional sentimental classification task, there are more changes in this task. Emotions are multidimensional, and each emotion has a degree. For example, the degree of happiness ranges from 0 to 5, with 0 being none and 5 being the strongest. A sentence may have a variety of emotions, such as joy, surprise. Emotion classification is for a certain role in a sentence, rather than the whole sentence. A sentence may have multiple roles with different emotions. Considering the property of the task, we tried a few networks which different from what the multi-classifier does.

数据：（做成表格，而不是图片，第一个情绪改成‘爱love’）





The above table is an example: The table content contains the script of a movie. The character column contains the specified character, that is mentioned in the script. The last 6 columns are the labels, which is in the training data but missing in the test data. The task is to identify the given character’s six emotions: love, happiness, surprise, anger, fear, and sorrow, and numerically rank them according to the script . A sentence has multiple characters, such as p2, d1 and x2, and for each character, the type and degree of emotion needs to be identified. In the sample, there is one line: A X2 Praise: “Wow, beautiful car!", which contains two emotions: "joy" and "surprise", and they are in degree 2 and 3, respectively.

带标签的数据总共有42896条，我们随机打乱后按8:2的比例分割成训练集和验证集。我们统计了在训练集上label分布，从数据分布中明显看出情绪值为0的占绝大多数。情绪值越大，占比越小。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| degree | love | joy | surprise | anger | fear | sorrow |
| 0' | 28434 | 27262 | 27735 | 26397 | 27048 | 24898 |
| 1' | 420 | 1645 | 1033 | 1612 | 1253 | 2259 |
| 2' | 328 | 370 | 458 | 981 | 815 | 1594 |
| 3' | 273 | 180 | 229 | 465 | 339 | 753 |
| percentage of 0 | 0.965337 | 0.925485 | 0.941606 | 0.896181 | 0.918282 | 0.843886 |
| percentage of 1 | 0.014259 | 0.055844 | 0.03507 | 0.054728 | 0.042539 | 0.076566 |
| percentage of 2 | 0.011136 | 0.012561 | 0.015549 | 0.033305 | 0.027669 | 0.054027 |
| percentage of 3 | 0.009268 | 0.006111 | 0.007775 | 0.015787 | 0.011509 | 0.025522 |

评估指标：

本任务采用常用的均方根误差（RMSE）来计算评分，按照“文本内容+角色名”识别出的6类情感对应的情感值来统计。

R M S E=\sqrt{\frac{\sum\_{i=1}^{n} \sum\_{j=1}^{6}\left(y\_{i, j}-x\_{i, j}\right)^{2}}{6 n}}*RMSE*=6*n*∑*i*=1*n*​∑*j*=16​(*yi*,*j*​−*xi*,*j*​)2​​

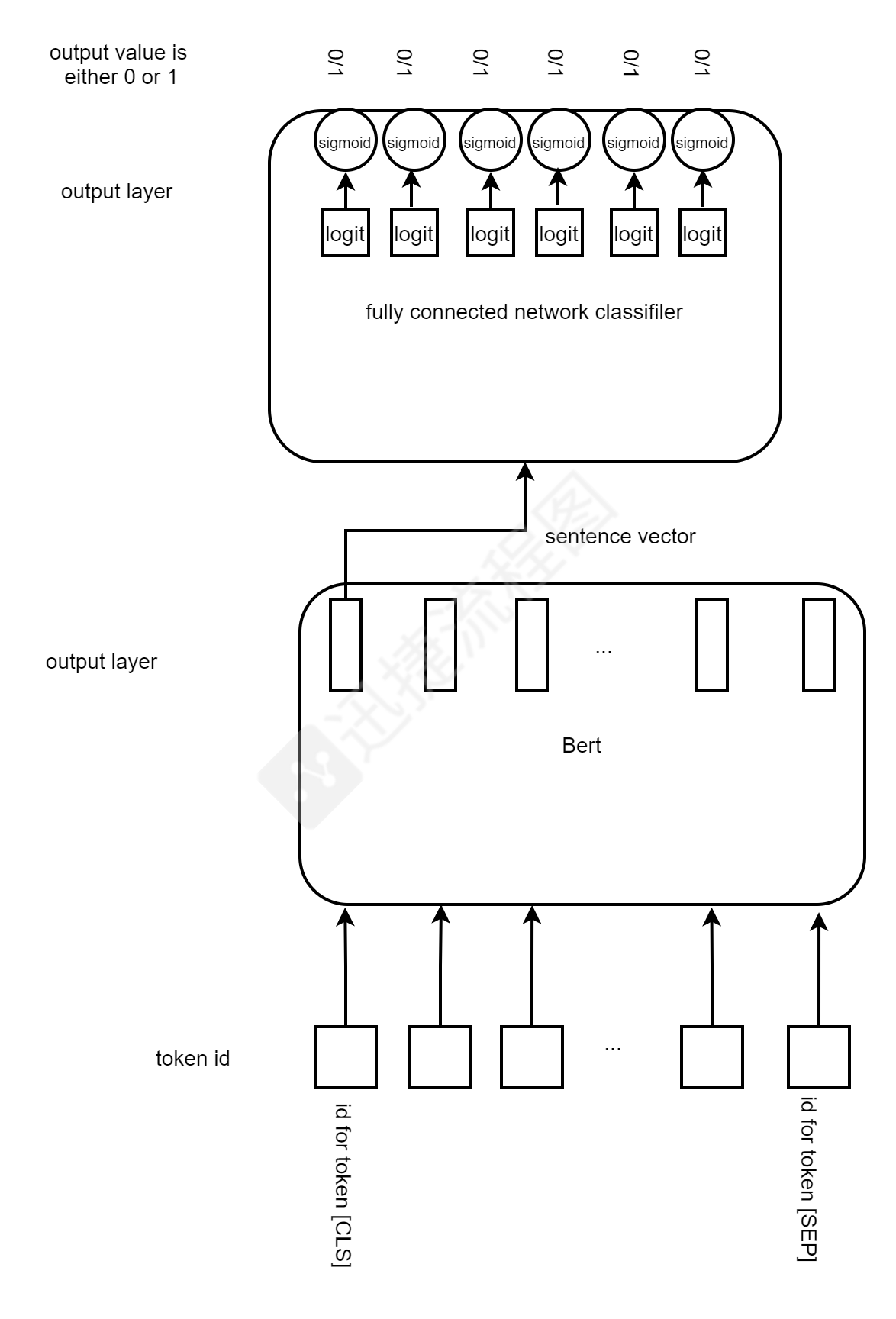
score = 1/(1 + RMSE)

其中是yi,j预测的情感值，xi,j是标注的情感值，n是总的测试样本数。  
**最终按score得分来排名。**

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



下一步的工作：

这个网络有个明显的缺点是，第一，把原本是多标签多分类的任务简化成多标签二分类的任务，这在一定程度降低了模型的性能。第二，要识别的角色仅仅是简单的与文本进行合并，没有明确告诉模型待识别的情绪到底是哪个角色。第三，角色的情绪值不仅仅依赖当前文本，有时还依赖历史文本，这个网络没有利用到历史信息进行预测。

我们下一步是尝试实现多标签多分类的网络。由于分类的类别是具有程度大小关系，我们将分类问题先当成回归问题，得到从0到3之间的实数值，然后再通过四舍五入的方式转成{0,1,2,3}中的某一个整数，从而得到分类结果。

再下一步，我们将参考问答任务，把角色当成query,与text进行attention,从而明确的告诉模型需要预测的哪个角色的情绪值。