报告

第一部分

模型：BGEM3FlagModel

参数：*batch\_size*=12, *max\_length*=100

混淆矩阵

图形用户界面, 文本

AI 生成的内容可能不正确。

准确率：80.8081%

使用阈值：0.83

阈值逼近：0.8112 – 0.8392

总结：

对于这个模型，估计是受限于模型大小，大部分出现TN and FP 都是因为近义词。对于模型来说，同义词（在这次实验中以微粒贷和贷款尤为突出）会极大的影响输出结果。解决办法：1. 使用一些库，使得给到模型的是已经做过预处理的数据，在这个步骤中消除近义词。2. 更换一个更大的模型

模型：bge-reranker-base

参数：use\_fp16=True 其余默认

混淆矩阵

表格

AI 生成的内容可能不正确。

使用阈值：0.74 – 0.84

阈值逼近：0.8276-0.8384

总结：这个模型对于中文的支持度更高。能把误差再降低10%。

阈值范围求到现在，建议使用的就是0.81和0.82，如果要求精度是二位浮点小数的话

第二部分

m3: 正例平均相似度 = 0, 负例平均相似度 = 0.47，差值：0.47

m3: 正例数量 = 0, 负例数量 = 315

bger: 正例平均相似度 = 0.93, 负例平均相似度 = 0.03，差值：0.89

bger: 正例数量 = 13, 负例数量 = 302

qwen: 正例平均相似度 = 0.83, 负例平均相似度 = 0.11，差值：0.71

qwen: 正例数量 = 2, 负例数量 = 313

(结果包含多位小数，已省略)