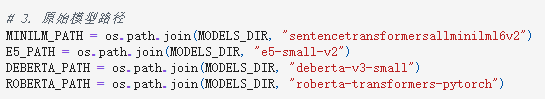
**Untitled.ipynb**

原始模型数据来源：  


最初模型：

模型 A LGBM (MiniLM + 4 特征) Validation LogLoss: 1.03534

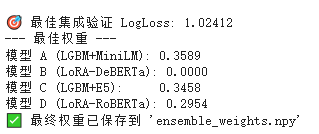
模型 B (LoRA DeBERTa-small) DeBERTa T = 1.084 | 校准后 Loss: 1.05883

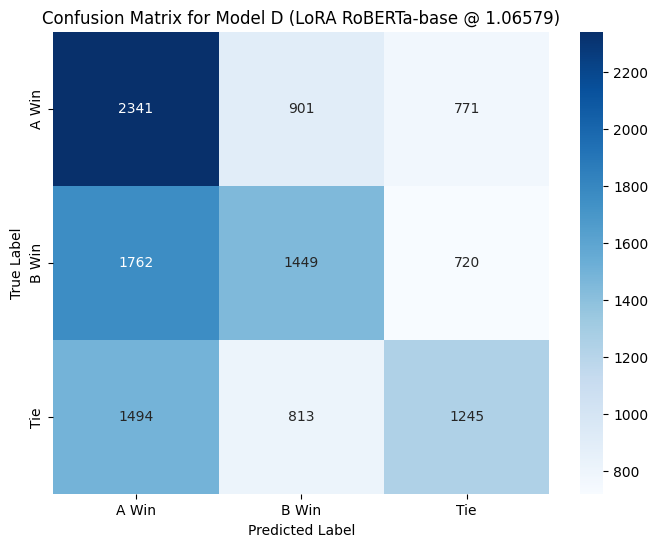
模型 C LGBM (E5 + 4 特征) Validation LogLoss: 1.03605

模型 D (LoRA RoBERTa-base) RoBERTa T = 1.244 | 校准后 Loss: 1.04997

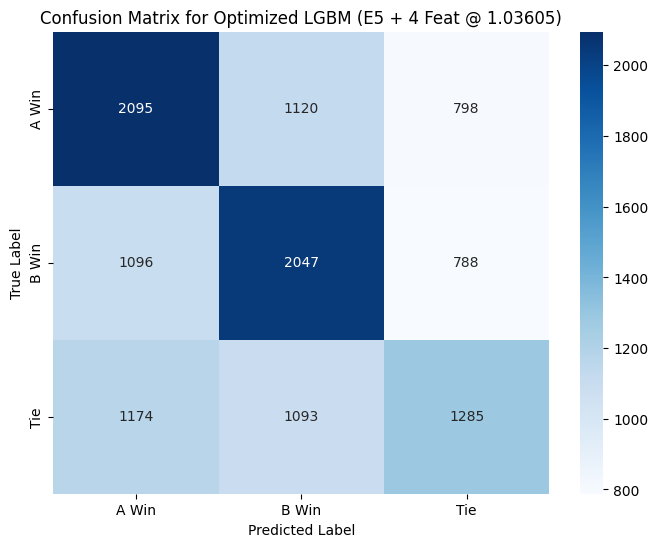
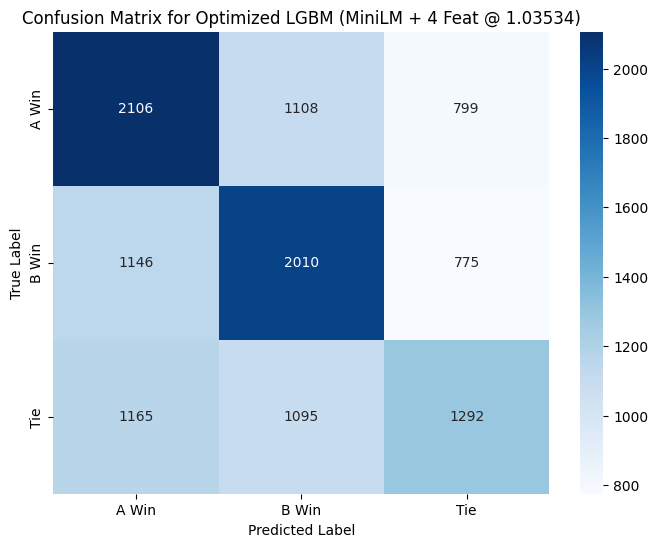
（其中ABC来自team8-model-extensions.ipynb，D是我新增的）

将ABCD四个模型集成：

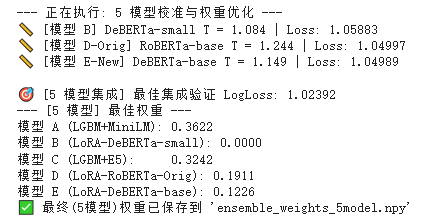


其中D的混淆矩阵如图，可以看出D对选择a win存在很大倾向，这被视为有价值的偏见，所以取得了0.29的权重  
、

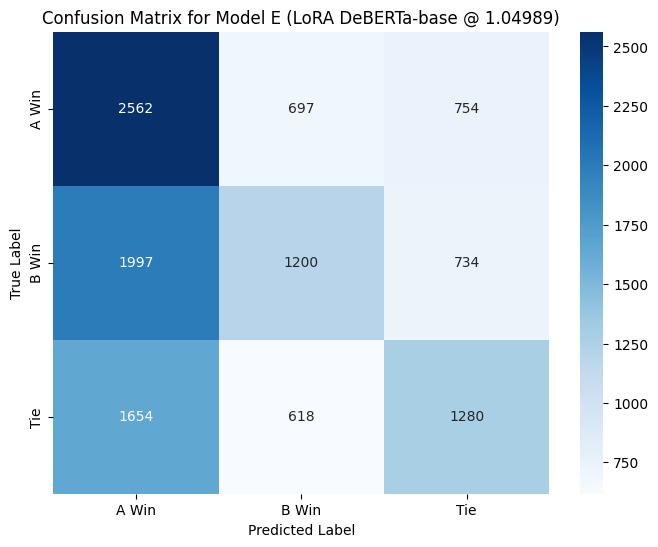
然后A和C的混淆矩阵如下图。可以看出两个模型非常相似，所以也取得了类似的权重（两个都是0.35左右）



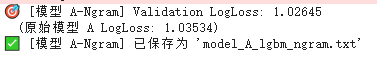
又由于模型B根本没有取得权重，猜测是因为DeBERTa-small数据集太小，所以新增了**模型 E (LoRA DeBERTa-v3-base)** Loss: 1.04989

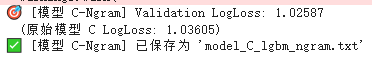


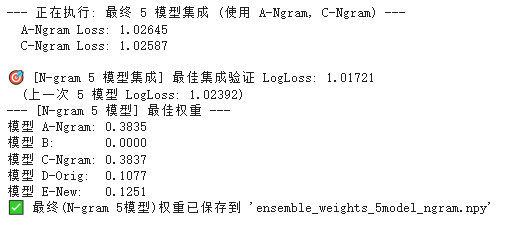
虽然集成后loss有提升但是非常小。

再进行对E的混淆矩阵分析。可以看出与D类似，都存在对A win的倾向。并且权重占比最大的还是A和C模型，所以考虑继续新增特征来优化权重占比最大的模型AC，节省时间成本  


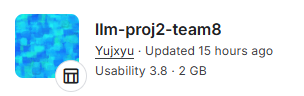
所以为LGBM 创建 N-gram 差异特征，重新训练了A和C







以上内容都在Untitled.ipynb文件内运行，保存训练后模型和参数信息，上传到kaggle内dataset，再次输出csv后提交成绩如图





**finalv4.ipynb**

在本地训练的流程中发现其实A和C模型之间没有太大区别，D和E之间也没有，并且由于D和E的训练时间太长，考虑直接进行简单训练

直接选择了模型A，使用MiniLM + 4 基础特征，加上之前从错误分析中新增的相似度特征和N-gram 差异特征，并且使用K折交叉验证（DE模型使用K折交叉验证时间花费太多）

得到最佳 [完整 OOF] 模型 A-Ngram **LogLoss：1.02799**

提交后成绩如图，取得了更好的成绩

