**模型训练使用特征:**

'gender\_final：候选人性别 Male/Female转化为1/0

'age'： 候选人年龄 由2024减去候选人出生年得到

'primary\_country',：候选人国籍 转化为类别

'english\_level', 'french\_level','arabic\_level', 'chinese\_level', 'russian\_level', 'spanish\_level'：

候选人语言能力 High/Intermediate/Low/No分别转化为 3/2/1/0

'lawflag', 'church\_flag', 'academic\_flag', 'firm\_flag','internationalorg\_flag', 'ngo\_flag', 'state\_flag'：

候选人背景信息 Yes/No转化为0/1

'academic\_current\_flag', 'internationalorg\_current\_flag','ngo\_current\_flag', 'state\_current\_flag',

'firm\_current\_flag', 'church\_current\_flag', 'la\_current\_flag'：

候选人最近一次工作类别 1/0

'mandate'：工作描述 转化为类别

'english\_level\_diff', 'french\_level\_diff', 'arabic\_level\_diff','chinese\_level\_diff', 'russian\_level\_diff',

'spanish\_level\_diff','lawflag\_diff', 'church\_flag\_diff', 'academic\_flag\_diff','firm\_flag\_diff',

'internationalorg\_flag\_diff', 'ngo\_flag\_diff','state\_flag\_diff', 'academic\_current\_flag\_diff',

'internationalorg\_current\_flag\_diff', 'ngo\_current\_flag\_diff','state\_current\_flag\_diff',

'firm\_current\_flag\_diff','church\_current\_flag\_diff', 'la\_current\_flag\_diff', 'age\_diff'：

候选人特征值与其面试组内平均特征值的差

**模型训练使用标签:**

**简历筛选轮：**

'是否进面试'：Yes/No转化为1/0

**面试轮：**

’selected‘：Yes/No转化为1/0

**训练测试划分：**

根据'session'和'工作描述'生成从0开始的整数作为每轮面试对应的ID，训练数据使用面试ID值前80%的数据，其余数据则作为测试数据用于计算准确率。此外，对于简历筛选轮，在划分数据之前剔除了存在全部候选人均进入面试的轮次。

**模型评估标准：**

对于测试数据，提取每轮面试对应的面试通过人数。接下来使用模型对数据进行预测，得到每个候选者通过面试的概率。接着，将每轮面试对应候选者的通过概率进行排序，并根据每轮面试的通过人数，选出概率从高到低相同数量的行作为预测通过，其余则预测为不通过。最后，通过比较测试数据原本的面试通过情况来计算预测的准确率。

**传统机器学习模型表现：**

**简历筛选轮：**

xgb = XGBClassifier(n\_estimators=36, learning\_rate=0.15, max\_depth=9,

subsample=1, min\_child\_weight=3, scale\_pos\_weight=1,

reg\_alpha=0, reg\_lambda=1, colsample\_bytree=1, gamma=0,

objective='binary:logistic', seed=1, enable\_categorical=True)

**准确率：**0.707808564231738

**面试轮：**  
xgb = XGBClassifier(n\_estimators=6, learning\_rate=0.1, max\_depth=2,

subsample=1, min\_child\_weight=3, scale\_pos\_weight=1,

reg\_alpha=0, reg\_lambda=1, colsample\_bytree=0.8, gamma=0,

objective='binary:logistic', seed=1, enable\_categorical=True)

**准确率：**0.7227722772277227

**DeepSeek大模型表现：**

由于大模型的输出存在不稳定性，将10次输出的准确率平均值作为最终的准确率。

**简历筛选轮：**

1. 无样例，准确率：0.665743073047859
2. 两个随机样例，准确率：0.6740554156171286
3. 有大模型根据全部训练数据总结得出的总结性文本，准确率：0.6586901763224182
4. 有大模型根据全部训练数据总结得出的总结性文本 + 两个随机样例，准确率：

0.6632241813602016

**面试轮：**

1. 无样例，准确率：0.6504065040650406
2. 两个随机样例，准确率：0.6504065040650406
3. 有大模型根据全部训练数据总结得出的总结性文本，准确率：0.6504065040650406
4. 有大模型根据全部训练数据总结得出的总结性文本 + 两个随机样例，准确率：

0.6178861788617886