* **模型使用数据:**

1. 职位名称：该组面试对应职位名称，如：Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance
2. 候选人个人信息：候选人国籍，年龄，性别
3. 候选人语言能力：根据候选人简历中语言能力得到，包括阿拉伯语/中文/英语/西班牙语/俄语/法语，将能力水平分类为：高/中/低/无
4. 候选人背景信息：根据候选人简历中工作和学历信息得到，包括是否有法律/宗教/公司/政府/学院/IGO/NGO方面的背景，分为：是/否
5. 候选人最近一份工作类型：根据候选人简历中最近一份工作内容得到，包括是否为法律/宗教/公司/政府/学院/IGO/NGO方面的类型，分为：是/否
6. 细节信息：候选人简历中具体教育经历及工作经历信息。对于传统机器学习，教育经历仅使用学位/专业/大学/大学所在国家，工作经历仅使用职位/公司/公司所在国家作为特征。

* **模型测试方法：**

**传统机器学习（XGBoost）：**

我们将每组面试的数据作为测试数据，将测试数据以外全部数据作为训练数据。对于测试数据，提取每组面试对应的面试通过人数，通过模型预测得到每个候选者通过的概率并将概率进行排序。根据每组面试的通过人数，选出概率从高到低相同数量的人数预测为通过，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及召回率。

在XGBoost的超参数设置情况上，我们在简历筛选轮中均使用默认参数，而在面试轮中，由于正负样本严重不均，我们使用官方文件的推荐方式设置scale\_pos\_weight，即负样本数/正样本数。

**大模型（DeepSeek）：**

由于数据中存在仅出现过一次的职位和出现过多次的职位，对于出现超过一次的职位信息我们将测试数据以外的相同职位数据作为该测试数据的对应样例。以下方法中对于仅出现过一次的职位均使用无样例方法生成的提示语，对于其余职位则使用该方法对应的提示语生成方法。最后根据大模型给出的通过候选人ID计算准确率和召回率。

1. 无样例：提示语仅包括测试数据，如：

You are a member of the UNHRC, based on the information of candidates, select who can be shortlisted for interview. Their mandate is Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance.

Please select "EXACTLY 6 candidates" using the candidates information below.

Give the ID numbers of the candidates that you have selected,do not explain why you have chosen the candidates nor rank them in order, just the ID numbers.Please respond with the following format: @@@ The candidates ID that you have selected are: @@@

The candidates information are:

Candidate ID: 1, She is a citizen of Russia. Her age is 63.0. She has high english level, high french level, low arabic level, low chinese level, high russian level, high spanish level. She has legal background, academic background, NGO background. Her current job is academic and law related.

Her educational background is as following: ......

Her working background is as following: ......

......

Remember you need to select EXACTLY 6 and only 6 candidates from the candidates information above.

1. 有样例：提示语在无样例的基础上，在测试数据之前添加对应样例数据以及通过候选人ID，如：

You are a member of the UNHRC, based on the information of candidates, select who can be shortlisted for interview. Their mandate is Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance.

Before selecting the candidates, this mandate has been mentioned in previous UNHRC meetings.

The candidates information in these meetings are:

Candidate ID: 1220, ......

In this case, the candidates that were shortlisted are: 1220, 1236, 1241, 1244, 1246.

Now that you have had some experience with this task,please select "EXACTLY 6 candidates" using the candidates information below.

Give the ID numbers of the candidates that you have selected,do not explain why you have chosen the candidates nor rank them in order, just the ID numbers.

Please respond with the following format: @@@ The candidates ID that you have selected are: @@@

The candidates information are:

Candidate ID: 1, ......

Remember you need to select EXACTLY 6 and only 6 candidates from the candidates information above.

1. 有总结文本：使用大模型将对应样例数据及通过情况进行总结作为指导添加进提示语，如：

You are a member of the UNHRC, based on the information of candidates, select who can be shortlisted for interview. Their mandate is Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance.

Before selecting the candidates, this mandate has been mentioned in previous UNHRC meetings.

Here is the commentary summary of the selected candidates in these previous UNHRC meetings:

The candidates selected for the role of Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia, and related intolerance possess a combination of legal expertise, academic credentials, and practical experience in human rights advocacy. Key attributes that make these candidates suitable include:

1. \*\*Substantial Legal and Academic Background\*\*: All candidates have advanced degrees in law, including Juris Doctor (JD) and PhDs, from prestigious institutions such as Yale Law School and the University of Michigan. This indicates a deep understanding of legal frameworks and human rights issues.

2. \*\*Experience in Human Rights Advocacy\*\*: Candidates have demonstrated experience in various capacities, including academic roles, legal practice, and positions within international organizations and NGOs. This experience spans teaching international human rights law, providing legal representation to survivors of human rights violations, and leading advocacy efforts against discrimination and racism.

3. \*\*International Exposure and Expertise\*\*: The candidates have worked or studied in multiple countries, reflecting a broad perspective on global issues related to racism and discrimination. Their international experience includes roles with the United Nations, where they have contributed to policy development and advocacy on human rights issues.

4. \*\*Policy and Advocacy Skills\*\*: Several candidates have held positions that required them to develop and implement policies aimed at combating racism and promoting equality. This includes coordinating efforts within government bodies and NGOs to address systemic issues related to discrimination.

5. \*\*Leadership and Management Skills\*\*: The candidates have shown leadership in various roles, from managing academic departments and leading NGOs to coordinating international human rights projects. Their ability to lead and manage diverse teams and initiatives is crucial for the role of Special Rapporteur.

These attributes collectively equip the candidates with the necessary skills and insights to effectively address and report on contemporary forms of racism, racial discrimination, xenophobia, and related intolerance at an international level.

Referring to this information, please select "EXACTLY 6 candidates" using the candidates information below.

Give the ID numbers of the candidates that you have selected,do not explain why you have chosen the candidates nor rank them in order, just the ID numbers.Please respond with the following format: @@@ The candidates ID that you have selected are: @@@

The candidates information are:

Candidate ID: 1, ......

Remember you need to select EXACTLY 6 and only 6 candidates from the candidates information above.

1. RuleSet: 使用深度为3的决策树学习对应样例数据并使用大模型总结出指导方案，如：

You are a member of the UNHRC, based on the information of candidates, select who can be shortlisted for interview. Their mandate is Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance.

From previous knowledge, we know that If a candidate is female, does not have a background in NGOs, and has intermediate or lower French language proficiency, the probability of being shortlisted is higher. Taking this as a reference (and only as a reference) for selecting the suitable candidates, please select "EXACTLY 6 candidates" using the candidates information below.

Give the ID numbers of the candidates that you have selected,do not explain why you have chosen the candidates nor rank them in order, just the ID numbers.Please respond with the following format: @@@ The candidates ID that you have selected are: @@@

The candidates information are:

Candidate ID: 1, ......

Remember you need to select EXACTLY 6 and only 6 candidates from the candidates information above.

1. Prototype: 使用K-Prototype方式（k=2)，将每个质心最近的类作为该质心对应的类，由大模型将每个质心信息进行提炼总结作为指导方案，如：

You are a member of the UNHRC, based on the information of candidates, select who can be shortlisted for interview. Their mandate is Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance.

From previous knowledge, we know that Generally, a successful candidate is associated with:

- Being in their late thirties, female, and proficient in English, while having no proficiency in Arabic, Chinese, French, Russian, or Spanish.

- Holding a background in academia and currently employed in an academic role.

- Having no affiliations with law, religious organizations, companies, international governmental organizations, NGOs, or government bodies, either in their background or current employment.

Conversely, a candidate less likely to be successful is associated with:

- Being in their mid-sixties, female, and proficient in English, with no proficiency in Arabic, Chinese, French, Russian, or Spanish.

- Having a background in both academia and government, but currently not employed in either field.

- Having no affiliations with law, religious organizations, companies, international governmental organizations, NGOs, or government bodies, except for a background in government.

Taking this as a reference (and only as a reference) for selecting the suitable candidates, please select "EXACTLY 1 candidates" using the candidates information below.

Give the ID numbers of the candidates that you have selected, do not explain why you have chosen the candidates nor rank them in order, just the ID numbers.

Please respond with the following format: @@@ The candidates ID that you have selected are: @@@

The candidates information are:

Candidate ID: 1, ......

Remember you need to select EXACTLY 6 and only 6 candidates from the candidates information above.

* **模型表现：**

**简历筛选轮：**

1. XGBoost（无细节信息） 准确率为：0.683 召回率为：0.521
2. XGBoost（有细节信息） 准确率为：0.687 召回率为：0.527
3. 大模型无样例（无细节信息） 准确率为：0.650 召回率为：0.472
4. 大模型有样例（无细节信息） 准确率为：0.639 召回率为：0.455
5. 大模型无样例（有细节信息） 准确率为：0.702 召回率为：0.550
6. 大模型有样例（有细节信息） 准确率为：0.704 召回率为：0.554
7. 大模型有总结文本（有细节信息） 准确率为：0.686 召回率为：0.526
8. 大模型RuleSet（有细节信息） 准确率为：0.660 召回率为：0.486
9. 大模型Prototype（有细节信息） 准确率为：0.665 召回率为：0.495

**面试轮：**

由于面试轮数据量过少，K-Prototype在许多组面试上无法计算，遂不包括Prototype方法。

1. XGBoost（无细节信息） 准确率为：0.689 召回率为：0.310
2. XGBoost（有细节信息） 准确率为：0.680 召回率为：0.289
3. 大模型无样例（无细节信息） 准确率为：0.651 召回率为：0.225
4. 大模型有样例（无细节信息） 准确率为：0.651 召回率为：0.225
5. 大模型无样例（有细节信息） 准确率为：0.660召回率为：0.246
6. 大模型有样例（有细节信息） 准确率为：0.687 召回率为：0.305
7. 大模型有总结文本（有细节信息） 准确率为：0.699 召回率为：0.332
8. 大模型RuleSet（有细节信息） 准确率为：0.670 召回率为：0.267

* **最新进度：**

我们发现在使用细节信息的情况下，大模型的表现严重依赖输入数据的顺序。为了解决这一问题，我们将候选人ID由数字转换为MD5编码以此避免大模型对特定ID有偏见。即便如此，在打乱原始数据顺序的情况下准确率总会有较为显著的下降，并且让大模型对于每个候选人提炼特点或给出选择原因均未能改善这一情况。同时如果使大模型单独对候选人打分，所有候选人的得分均会在2-3个偏高得分中变动。

为了解决顺序依赖的问题，我们改为对于每组面试多次获取大模型对于不同顺序状态下的回答并将被选中次数最多的那些候选人作为大模型的预测，结果如下：

**简历筛选轮：**

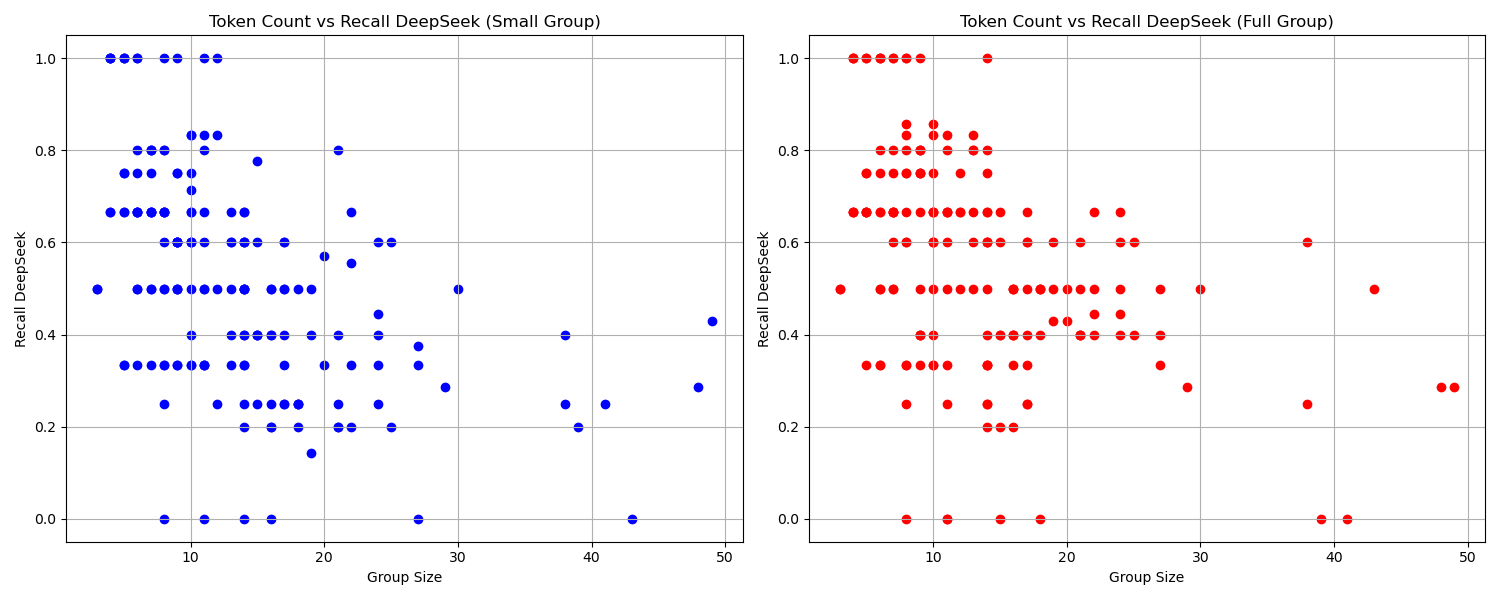
1. 无样例（10次投票） 准确率为：0.685 召回率为：0.524
2. 有样例（10次投票） 准确率为：0.683 召回率为：0.519
3. 总结文本（10次投票） 准确率为：0.696 召回率为：0.541
4. 无样例（20次投票） 准确率为：0.704 召回率为：0.552

发现在多次投票的情况下，大模型的表现有明显回暖，其中使用大模型总结文本的方法有着相对最稳定的结果。

另一种猜测是过长的提示语或过多的候选人会导致大模型表现下降，为了验证这一猜测，我们将原本使用所有候选人让大模型进行预测的方法改为对于每组面试，每次仅给大模型 待选人次+1 个候选人，然后淘汰掉未被选中的那个候选人，不断重复这一过程，最终确定候选人名单。然而这一方法的表现并不突出，结果为：

**简历筛选轮：**

1. 无样例 准确率为：0.678 召回率为：0.515

通过下面的散点图也可以看到使用小分组情况下召回率和组内人数的变化情况与将全部候选人同时进行预测并无太大区别，说明大模型表现似乎并不与候选人数相关。

1. 有样例 准确率为：0.68 召回率为：0.514
2. 总结文本 准确率为：0.68 召回率为：0.516