

# 第十七届“华中杯”大学生数学建模挑战赛题目

---

## C 题 就业状态分析与预测

就业是最基本的民生，是经济发展的重要支撑。当前，我国就业形势保持基本稳定，但也面临一些挑战，就业结构性矛盾尚存在。促进高质量充分就业，是宏观经济政策的重要目标之一，是新时代新征程就业工作的新定位、新使命。高质量充分就业常受到宏观经济政策、经济发展水平等诸多因素影响，精准的就业状态预测对充分就业政策制订具有重要参考价值。本赛题以宜昌地区部分就业数据为研究对象，旨在通过数学建模方法，对就业状态进行分析和预测，进而为相关部门制定符合当地实际的就业政策提供科学的决策依据。

本赛题提供了宜昌地区 5000 名被调查者的脱敏数据（见附件 1），其中包含 53 个变量，涵盖被调查者的个人基本信息（比如：年龄、民族、文化程度等 29 个变量）、就业信息（比如：就业时间、行业代码、录用单位等 9 个变量）、失业信息（比如：失业时间、失业原因、原从事工种等 15 个变量）以及预测集 20 个样本。

根据上述数据，请完成以下任务：

### 问题一：数据特征分析

请参赛者根据被调查者当前的就业状态（比如：就业失业时间、录用单位等信息）分析该地区当前就业的整体情况；并将人员按照年龄、性别、学历、专业、行业等特征进行划分，根据划分特征分析其对就业状态的影响。

要求：在作品中明确对就业状态进行标注，请以示例表 1 格式给出就业失业状态数量；对不同层面因素的影响用图表形式进行展示。

示例 表 1 当前就业状态

| 就业失业状态 | 就业 | 失业 |
|--------|----|----|
| 数量（人）  |    |    |

问题二：就业状态预测

基于问题一的分析，选取与就业状态具有相关性的特征，构建就业状态预测模型并对附件 1 中给定的“预测集”进行预测；并对各特征的重要性进行排序。

要求：在作品中使用准确率、查准率、召回率、F1 等指标对模型进行评估，并用示例表 2 格式给出各评估指标的结果，用条形图绘制重要特征的排序，以示例表 3 格式给出其预测结果。

示例 表 2 评价指标结果

| 模型      | 准确率 | 查准率 | 召回率 | F1 |
|---------|-----|-----|-----|----|
| Model-1 |     |     |     |    |
| .....   |     |     |     |    |

示例 表 3 就业状态预测结果

| 预测   | T1 | ..... | T20 | 就业数量<br>小计 | 失业数量<br>小计 |
|------|----|-------|-----|------------|------------|
| 就业状态 | 1  |       | 0   |            |            |

注：1 表示就业，0 表示失业

问题三：就业状态预测模型优化

除了个人层面因素影响外，宏观经济、政策、劳动力市场状况、宜昌市居民消费价格指数、招聘信息等也可能会影响就业状态。请参赛者收集相关数据，提取反映经济、市场等方面的影响因素，并结合问题一中的数据进一步完善就业状

态预测模型，并对附件 1 中给定的“预测集”进行预测。

要求：在作品中以表格形式给出你选择的外部变量和数据来源；使用准确率、查准率、召回率、F1 等指标进行评估，并用表格形式给出各评估指标的计算结果。

#### **问题四：人岗精准匹配**

基于赛题提供的数据，并结合采集到的招聘数据、社交媒体数据、薪资水平、所需技能、宏观经济数据、行业动态数据等建立人岗匹配模型，捕捉求职者和岗位之间的匹配关系，针对赛题数据中的失业人员进行工作推荐。

要求：在作品中以表格形式给出所考虑的外部变量和来源。

注：获取的外部数据、代码，请以附件提交。