# 人工智能-第三次课程作业说明

1. **问题描述**

请利用决策树判断是否需要在餐馆等座位，本问题主要基于如下属性列表基础上决策的。

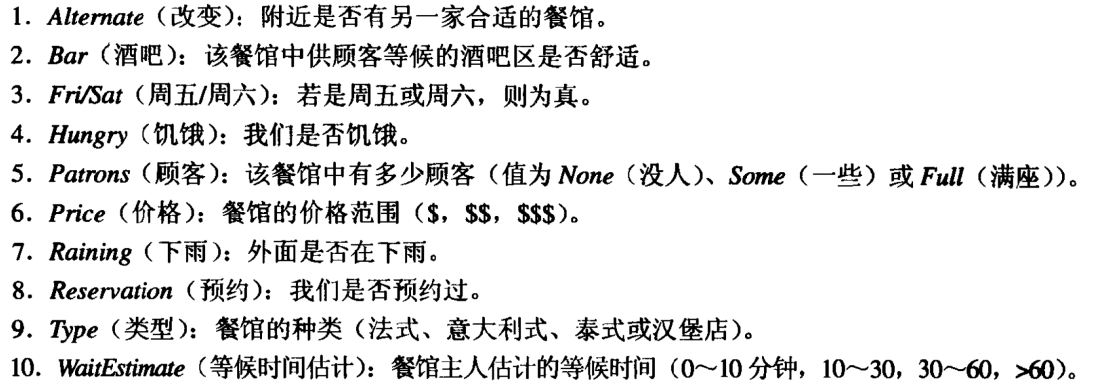


图1 餐馆等座位问题属性列表

1. **数据集介绍**

restaurant\_willwait数据集共有100条数据，按照3:1划分比例分为训练集和测试集，现提供训练集数据，即75条数据。保存路径为“Assignment3\_Refactor\dataset”大家可根据需要自行分为训练集和测试集。为了方便程序编写，数据集属性值取值做了如下对应改写：

表1 属性值离散表示

|  |  |
| --- | --- |
| 属性 | 离散取值 |
| Alternate | 1 0 |
| Bar | 1 0 |
| Fri/sat | 1 0 |
| Hungry | 1 0 |
| Patrons | 0 0.5 1 |
| Price | 0 0.5 1 |
| Raining | 1 0 |
| Reservation | 1 0 |
| Type | 1 0.67 0.33 0 |
| Waitestimate | 0 0.17 0.5 1 |

表2 是否等待离散表示

|  |  |
| --- | --- |
| 属性 | 离散取值 |
| willwait | 1 0 |

1. **项目说明**

使用VS2019打开Assignment3.sln文件，main.cpp中的main函数为项目的主入口。作业完成方式：（1）需要框架帮助的同学可以直接完成DecisionTree.cpp文件中的相应函数。（2）想自己构建框架的同学们需要自己构建决策树以及决策树中用到的相关函数，最后提供一个接口函数。接口函数的输入为样例列表，输出为整数列表（vector<int>）。

作业评估时，我们会调用该函数，把结果保存到classify\_res中，然后通过util中的evaluate函数评估准确性。

最终评测将在测试集上采用**正确率(A, Accuracy)**指标进行评测。

**[1]**

1. **作业要求**

**4.1文档部分（2分）**

* 针对本题目撰写文档，保证文档结构完整性，文档可包括 题目、问题描述、实验结果、分析、结论，也可参考其他标准文档或者学术论文的格式。
* 文档内容及结构均纳入评分范围。
* **文档：1分。源代码风格：1分**

**4.2源代码部分（8分）**

* 代码包括决策树学习和信息增益计算，最终评分按照决策树在测试集上的表现为准。

表3 评分标准

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 正确率 | 60%<70% | 70%A<80% | 80%A<90% | A90% |
| 分数 | 2 | 4 | 6 | 8 |

* 评价：希望同学们独立完成。有困难可以和其他同学讨论，学会思路，不要抄袭。

1. **作业提交**

提交时间：第12周周四（11月24日）上课前

提交方式：将电子版源代码及文档打包，命名为“学号\_姓名\_Assignment3”，发送至邮箱tengh@seu.edu.cn。

注意：打包方式使用zip；打包时不要把项目中的.vs文件打包进去。