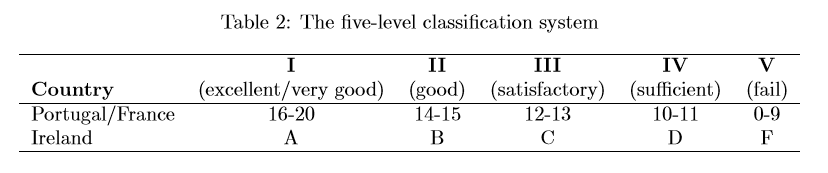
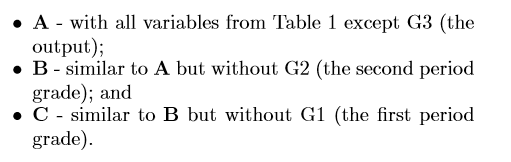
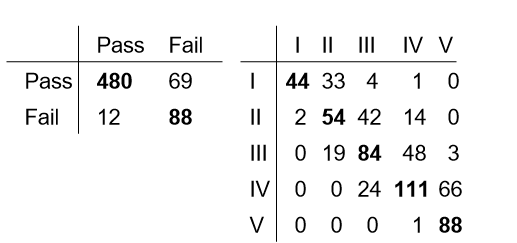
**机器学习实验**

1. 仔细阅读文件“数据属性说明.txt”，熟悉数据的格式及内容；
2. 两种途径：第一，weka平台，将数据导入Weka，并用arff格式存贮；（熟悉软件，会用）第二，python，利用sklearn机器学习包（推荐第二种方式，建议用python实现）
3. 对学生的成绩进行预测，二分类binary classiﬁcation (pass/fail)

pass if G3≥10, else fail

1. 对学生的成绩进行预测，进行5等级分类classiﬁcation with ﬁve levels (from I very good or excellent to V - insuﬃcient)



1. 用线性回归对学生的成绩进行预测。
2. 利用不同的属性集A、B、C进行实验，比较三种属性集下的准确率。
3. 选择一次实验结果，给出混淆矩阵，如图所示。
4. 将实验过程及结果记录，写成实验报告。
5. 扩展问题（可选）
6. 对于这个数据集，你还能做什么分析？
7. 使用sklearn的决策树模块时，如果数据是离散型取值，该怎么处理？